Universidad Autónoma de Madrid - UAM

Departamento de Bioquímica

DOCTORAL THESIS

NEW INSIGHTS IN MYELOID MALIGNANCIES: APPLICATION OF MASSIVE GENOMICS APPROACHES FOR DIAGNOSIS AND MONITORING

Juliane Menezes

Madrid, 2013

DEPARTAMENTO DE BIOQUÍMICA
FACULTAD DE MEDICINA
UNIVERSIDAD AUTÓNOMA DE MADRID - UAM

New insights in myeloid malignancies: application of massive genomics approaches for diagnosis and monitoring

Doctoral thesis submitted for the degree of Doctor in Biochemistry, Molecular Biology, Biomedicine and Biotechnology at Universidad Autónoma de Madrid by MSc in Biology

Juliane Menezes

Thesis Directors:

Dr. Juan Cruz Cigudosa

Dra. Sara Álvarez

GRUPO DE CITOGENÉTICA MOLECULAR
PROGRAMA DE GENÉTICA DEL CANCER HUMANO
CENTRO NACIONAL DE INVESTIGACIONES ONCOLÓGICAS - CNIO

Dr. Juan C. Cigudosa García, Jefe del Grupo de Citogenética Molecular del Programa de Genética del Cáncer Humano del Centro Nacional de Investigaciones Oncológicas (CNIO) y como Director de la Tesis, y

Dra. Sara Álvarez de Andrés, Investigadora del Grupo de Citogenética Molecular del Programa de Genética del Cáncer Humano del Centro Nacional de Investigaciones Oncológicas (CNIO) y como Directora de la Tesis,

CERTIFICAN:

Que Doña **Juliane Menezes** ha realizado el presente trabajo: **"New insights in myeloid malignancies: application of massive genomics approaches for diagnosis and monitoring"** y que a su juicio reúne plenamente todos los requisitos necesarios para optar al **grado de Doctor**, a cuyos efectos será presentado en la Universidad Autónoma de Madrid. El trabajo ha sido realizado bajo su dirección, autorizando su presentación ante el Tribunal Calificador.

Y para que se conste se extiende el presente certificado,

Madrid, Septiembre de 2013

VºBº del Director de la Tesis:

Juan C. Cigudosa

Sara Álvarez

UNIVERSIDAD AUTONOMA
DE MADRID

Dr. Leandro Sastre, Investigador del Instituto de Investigaciones Biomédicas de Madrid, CSIC/UAM, y como la Tutor de la Tesis,

CERTIFICA:

Que Doña **Juliane Menezes** ha realizado el presente trabajo: **"New insights in myeloid malignancies: application of massive genomics approaches for diagnosis and monitoring"** y que a su juicio reúne plenamente todos los requisitos necesarios para optar al **grado de Doctor**, a cuyos efectos será presentado en la Universidad Autónoma de Madrid. El trabajo ha sido realizado bajo su supervisión, autorizando su presentación ante el Tribunal Calificador.

Y para que se conste se extiende el presente certificado,
Madrid, Septiembre de 2013

VºBº del Tutor de la Tesis:

Leandro Sastre

This thesis has been performed in the Molecular Cytogenetics Group at Spanish National Cancer Research Center – CNIO from November 2009 to September 2013. The work was done under Dr. Juan C. Cigudosa and Dr. Sara Álvarez supervision and supported by an INTRASALUD project (PI 12-00425) and Obra Social- Fundación “La Caixa” fellowship to J.M. (G.02609).

NEW INSIGHTS IN MYELOID MALIGNANCIES:

APPLICATION OF MASSIVE GENOMICS APPROACHES FOR DIAGNOSIS AND MONITORING

“Research is to see what everybody else has seen and to think what nobody else has thought”

~ Albert Szent-Gyorgyi (1937 Nobel Prize for Medicine, 1893-1986)

À minha **querida mãe** por me ensinar que com fé e perseverança posso alcançar todos os meus objetivos.

ACKNOWLEDGMENTS
AGRADECIMENTOS

First of all I would like to thank my thesis director **Dr. Juan Cruz Cigudosa** for his enormous support and patience during this time that I have spent in his laboratory: You were always there to help me with advice and motivation.

Moreover I would like to acknowledge the members of my **PhD thesis committee**, Dr. Sara Álvarez, Dr. Mercedes Robledo and Dr. Mª José Calasanz for very important feedback during our meetings.

I am also very grateful to all members (current and former) of **Molecular Cytogenetics Group**: Alba Maiques, Almudena Gil, Álvaro Eguileor, Álvaro Wandosell, Amanda Sánchez, Ana del Rio, Bibiana Ferreira, Carlos Benítez, Francesco Acquadro, Glória Soler, Lorena Rollón, Luis Espinosa, Margarita Sánchez Beato, Mª Carmen Martín, Mirian Hernando, Sandra Rodríguez and Sonia Sánchez Guillén. My special thanks to **Carmen Carralero** and **Miguel Ángel Grillo** for their friendship (it would not be possible without you); and to **Rocío Salgado** and **Sara Álvarez**, who had the patience to share their knowledge and experiences.

To all other people from **CNIO** (that are too many to mention) who helped me somehow at work or simply shared with me the difficulties of everyday: all members of **Human Cancer Genetics Programme** (especially to Alicia Barroso, Maika González, Fátima Mercadillo, Agnieszka Maliszewska, Marta Kamieniak and André Amaral), **Bioinformatics Core Unit** (David Pisano, Gonzalo Gómez and Osvaldo Graña) and **Molecular Diagnostics Clinical Research Unit** (Luis Lombardia, Silvia Mota, Julio César Corral and Karin Tang).

Aos meus amigos do Brasil pela força, palavra amiga, alegria e toda a ajuda. Mesmo com um oceano no meio e nos encontrando uma vez por ano sabemos que amizade como a nossa é para sempre: **Ana Paula Fontão**, **Bárbara Alemar**, **Beatriz Vasconcello, Isabela Alvim, Jeniffer Dantas, Luciana Moreira, Kátia Candian, Ricardo Ramirez, Sophie Herszterg** e tantos outros amigos que passaram pela minha vida.

A todos os meus **familiares** (família Menezes e família Braga Areal). Todos vocês, cada um de uma forma peculiar, iluminaram meu caminho.

Ao **Rômulo Braga Areal**, por todos os momentos felizes que passamos juntos e todos os que virão nessa nova etapa de nossas vidas: Não tenho palavras para agradecer tudo o que você fez e faz por mim, por isso vou resumir em duas: Te amo.

À minha **mãe** (Lucia Menezes), meus **irmãos** (Flavia Menezes e Jarek Sochacki) e **Wilma** por todo o carinho, paciência, amor incondicional e tantas outras coisas que somente nossa família pode nos proporcionar.

A **ELE**, inteligência e causa suprema de todas as coisas, agradeço acima de tudo pela benção da vida e pelas oportunidades que tenho.

ABSTRACT
RESUMEN

Myeloid malignancies are clonal diseases originated of genetic alterations in hematopoietic stem or progenitor cells. Application of next-generation sequencing to myeloid leukemia has already yielded important discoveries, including the identification of common gene mutations (eg, *IDH1* and *DNTM3A*). However, these efforts have so far excluded rare myeloid disorders. In this work we used a systematic genomic strategy based on whole-exome sequencing, RNA sequencing and target re-sequencing approaches to provide deeper insights into the genetic pathogenesis of blastic transformation of chronic myeloid leukemia, chronic neutrophilic leukemia and blastic plasmacytoid dendritic cell neoplasm.

We identified mutations responsible for these diseases that occur in several genes whose encoded proteins belong principally to five classes: signaling pathways proteins (e.g. *FLT3*, *RAS*), transcription factors (e.g. *IKZF* family, *RUNX1*, *ZEB2*), epigenetic regulators (e.g. *ASXL1*, *DNMT3A*, *IDH1*, *IDH2*, *TET2*), tumor suppressors (e.g. *TP53*), and components of the spliceosome (e.g. *LUC7L2*, *SF3B1*, *U2AF1*). First, our data suggest that sequencing a wider panel of genes that includes *ASXL1*, *TP53* and *IKZF* gene family could be beneficial in the clinical management of non-responders chronic myeloid leukemia patients, while current diagnostic procedures recommend the study of *ABL1* mutations. Second, we show the concurrence of several genetic mechanisms that cooperate with the *CSF3R* mutation in chronic neutrophilic leukemia, such as *U2AF1*, *TET2*, *LUC7L2* and *ASXL1* mutations, an altered pattern of splicing in *RUNX1* gene and an expressed fusion gene *PIM3-SCO2*. Third, we found that dendritic cell leukemia has a mutational profile strikingly similar to that of other well-defined myeloid leukemias and patients with mutations in genes in the methylation pathways had a significantly reduced cumulative survival.

In conclusion, large-scale sequencing efforts allowed us the establishment of a comprehensive repertoire of these mutations, contributing to an improvement in the definition and classification of these myeloid malignancies, and to the identification of new prognostic markers and therapeutic targets.

Las neoplasias malignas mieloides son enfermedades clonales derivadas de alteraciones genéticas en células madre hematopoyéticas. La aplicación de las técnicas secuenciación de última generación en el estudio la leucemia mieloide ya ha producido importantes descubrimientos como la identificación de mutaciones genéticas comunes. Sin embargo, estos esfuerzos han excluido, hasta ahora, los trastornos mieloides raros. En este trabajo utilizamos una sistemática estrategia genómica basada en la secuenciación masiva del exoma, del ARN total y de genes diana seleccionados (re-secuenciación), con el objetivo de proporcionar una mejor comprensión de la patogénesis genética de la transformación blástica en la leucemia mieloide crónica, la leucemia neutrofílica crónica y la neoplasia de células dendríticas.

Se identificaron mutaciones responsables de estas enfermedades que se producen en genes cuyos productos proteicos pertenecen principalmente a cinco clases de proteínas: vías de señalización (*FLT3*, *RAS*), factores de transcripción (*IKZF* familia, *RUNX1*, *ZEB2*), reguladores epigenéticos (*ASXL1*, *DNMT3A*, *IDH1*, *IDH2*, *TET2*), supresores tumorales (*TP53*), y componentes de la maquinaria de *splicing* (*LUC7L2*, *SF3B1*, *U2AF1*). En primer lugar, nuestros datos sugieren que la secuenciación de un panel más amplio de genes que incluya *ASXL1*, *TP53* y la familia de genes *IKZF* podría ser beneficiosa para el manejo clínico de los pacientes con leucemia mieloide crónica no respondedores. En segundo lugar, se presenta la concurrencia de varios mecanismos genéticos, que cooperan con la mutación *CSF3R* en la leucemia neutrofílica crónica, como las mutaciones en *U2AF1*, *TET2*, *LUC7L2* y *ASXL1*, un modelo diferente de procesamiento del ARN en el gen *RUNX1* y un gen de fusión *PIM3-SCO2*. En tercer lugar, hemos encontrado que la leucemia de células dendríticas tiene un perfil mutacional sorprendentemente similar al de otras leucemias mieloides bien definidas y que los pacientes con mutaciones en los genes en las vías de metilación presentaban una supervivencia acumulada significativamente reducida.

En conclusión, la secuenciación masiva de enfermedades mieloides raras nos permitió el establecimiento de un repertorio completo de estas mutaciones, lo que contribuye a una mejora en la definición y la clasificación de estas enfermedades, así como a la identificación de nuevos marcadores de pronóstico y de dianas terapéuticas.

ABBREVIATIONS

ABL1	c-abl oncogene1
AML	acute myeloid leukemia
AP	accelerated phase
APL	acute promyelocytic leukema
ASXL1	additional sex combs like 1
BC	blast crisis
BCR	breakpoint cluster region
BM	bone marrow
bp	base pairs
BPDCN	Blastic plasmacytoid dendritic cell neoplasm
CCyR	complete cytogenetic response
CDKN2A	cyclin-dependent kinase inhibitor 2A
cDNA	complementary DNA
CGH	comparative genome hybridization
ChIP	chromatin immunoprecipitation
CHR	complete hematological remission
CML	chronic myeloid leukemia
CMML	chronic myelomonocytic leukemia
CNL	chronic neutrophilic leukemia
CP	chronic phase
CSF3R	colony stimulating factor 3 receptor
del	deletion
DNA	deoxyribonucleic acid
DNMT3A	DNA methyltransferase 3A
ECOG	Eastern Cooperative Oncology Group
FISH	fluorescent in situ hybridization
FLT3	fms-related tyrosine kinase 3
gDNA	genomic DNA
HOXB9	homeobox B9
IDH1	isocitrate dehydrogenase 1
IDH2	isocitrate dehydrogenase 2
IKZF1	IKAROS family zinc finger 1
IKZF3	*IKAROS* family zinc finger 3 - Aiolos
indels	small insertions and deletions
inv	invertion
Kb	kilobases
LAP	leukocyte alkaline phosphatase
LOH	loss of heterozygozity
LUC7L2	LUC7-like 2
MB	megabases
MDS	myelodysplastic syndrome
MPN	myeloproliferative neoplasm
mRNA	messenger RNA
NCgR	no cytogenetics response
NGS	next-generation sequencing
NPM1	nucleophosmin 1
OS	overall survival
p	short arm of a chromossome
PB	peripheral blood
PCR	polymerase chain reaction
Ph	Philadelphia chromosome
q	long arm of a chromossome
RB1	retinoblastoma 1
RNA	ribonucleic acid
rob	Robertsonian translocations
RT-PCR	reverse transcriptase–polymerase chain reaction
RUNX1	runt-related transcription factor 1
SNP	single nucleotide polymorphism
SNS	single nucleotide substitutions
SNV	single nucleotide variants
t	translocation
TET1	tet methylcytosine dioxygenase 1
TET2	tet methylcytosine dioxygenase 2
TKI	tyrosine kinase inhibitors
TP53	tumor protein p53
U2AF1	U2 small nuclear RNA auxiliary factor 1
UBE2G2	ubiquitin-conjugating enzyme E2G 2
UPD	uniparenteral disomy
UTR	untranscribed regions
WBC	white blood cells
WES	whole-exome sequencing
WGS	whole-genome sequencing
WHO	world health organization
ZEB2	zinc finger E-box binding homeobox 2

TABLES

FIGURES

CONTENTS

Contents

CHAPTER 1:
Introduction

MYELOID MALIGNANCIES are clonal disorders of hematopoietic stem cells, in which there is an increased proliferation of the myeloid series, which leads to peripheral blood leukocytosis, increased erythrocyte mass or thrombocytosis [1]. This heterogeneous group of diseases is divided into five main categories, according to the 2008 world health organization (WHO): (1) myeloproliferative neoplasms; (2) myeloid neoplasms associated with eosinophilia and abnormalities of PDGFRA, PDGFRB, or FGFR1; (3) myelodisplastic/myeloproliferative neoplasms; (4) myelodisplastic syndrome; and (5) acute myeloid leukemia and related neoplasms (Table 1). The WHO classification criteria as well the description of the subtypes of myeloid malignancies that were studied in this doctoral thesis are detailed in the following sections.

1.1 The 2008 World Health Organization (WHO) classification of myeloid malignancies

The WHO classification uses all available information – morphology, cytochemistry, immunophenotype, genetics, and clinical features – to define clinically significant disease entities. It is a consensus classification in which a number of experts have agreed on the classification and the diagnostic criteria used for defining the entities that compose it.

In the WHO classification, the term “myeloid” includes all cells belonging to the granulocytic (neutrophil, eosinophil, basophil), monocytic / macrophage, erythroid, megakaryocytic and mast cell lineages (Figure 1). The WHO criteria for myeloid neoplasms apply to initial diagnostic peripheral blood (PB) and bone marrow (BM) specimens obtained prior to any definitive therapy for a suspected hematologic neoplasm. Morphologic, cytochemical and/or immunophenotypic features are used for establishing the lineage of the neoplastic cells and for assessment of their maturation. The blast percentage remains a practical tool for categorizing myeloid neoplasms and judging their progression. In the WHO scheme, a myeloid neoplasm with 20% or more blasts in the PB or BM is considered to be acute myeloid leukemia (AML) when it occurs *de novo*, evolution to AML when it occurs in the setting of a previously diagnosed myelodysplastic syndrome (MDS) or myelodysplastic/ myeloproliferative

neoplasm (MDS/MPN), or blast transformation in a previously diagnosed myeloproliferative neoplasm (MPN).

A complete cytogenetic analysis of BM cells is essential during initial evaluation for establishing a baseline karyotype; repeat analyses are recommended as needed thereafter for judging the response to therapy or for detecting genetic evolution. Additional genetic studies should be guided by the results of the initial karyotype and by the diagnosis suspected based on the clinical, morphologic, and immunophenotypic studies. In some cases, reverse transcriptase–polymerase chain reaction (RT-PCR) and/or fluorescence in situ hybridization (FISH) may detect variants of well-recognized cytogenetic abnormalities or submicroscopic abnormalities not detected by routine karyotyping, such as the *FIP1L1-PDGFRA* fusion in some myeloid neoplasms associated with eosinophilia [2]. In addition, gene mutations are increasingly being recognized as important diagnostic and prognostic markers in myeloid neoplasms. These include, among others, mutations of *JAK2*, *MPL*, and *KIT* in MPN [3-8]; *NRAS*, *KRAS*, *NF1*, and *PTPN11* in MDS/MPN [9-14]; *NPM1*, *CEBPA*, *FLT3*, *RUNX1*, *KIT*, *WT1*, and *MLL* in AML [15-19]; and *GATA1* [20] in myeloid proliferations associated with Down syndrome. Many of these gene mutations figure importantly in the revised WHO classification.

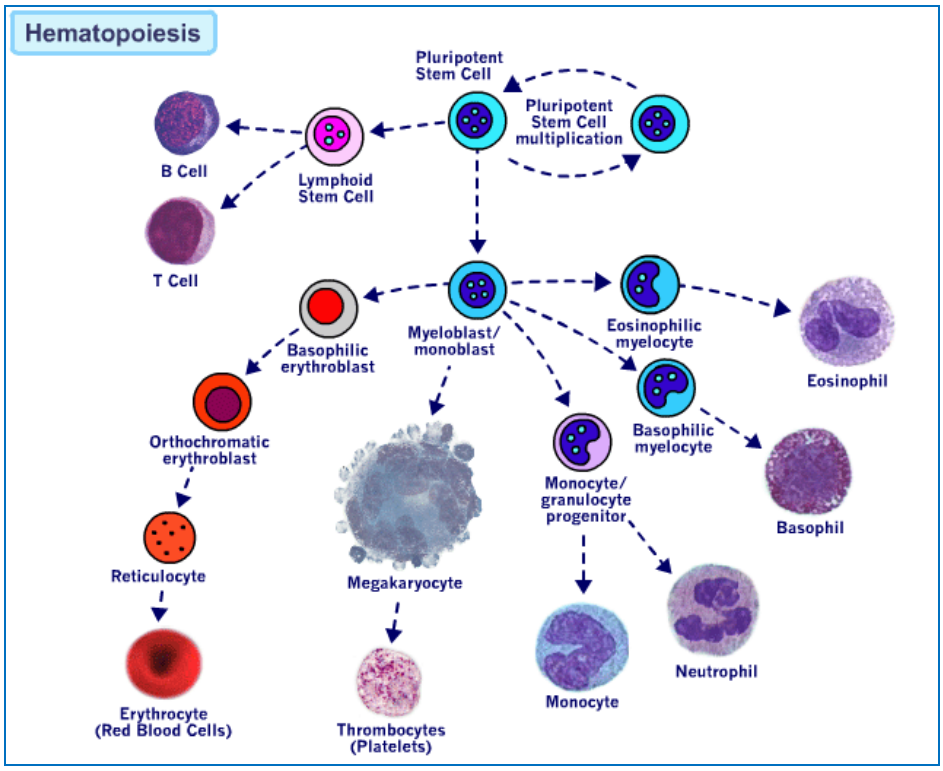

Figure 1: Classical hierarchal map of myelopoiesis: formation of myeloid cells **from** the pluripotent hematopoietic stem cells in the bone marrow via myeloid stem cells. Myelopoiesis generally refers to the production of leukocytes in blood, such as monocytes and granulocytes. This process also produces precursor cells for macrophage and dendritic cells found in the lymphoid tissue (© McGill Molson Medical Informatics Project).

Table 1 lists the major subgroups of myeloid malignancies in the WHO classification, and the specific entities of which they are composed. The blue ones are entities that were studied in this thesis.

Table 1: WHO classification of myeloid neoplasms (adapted from Vardiman *et. al.* 2009 [21])

- **Myeloproliferative neoplasms (MPN)**
  - Chronic myelogenous leukemia, *BCR-ABL1*-positive
  - Chronic neutrophilic leukemia
  - Polycythemia vera
  - Primary myelofibrosis
  - Essential thrombocythemia
  - Chronic eosinophilic leukemia, not otherwise specified
  - Mastocytosis
  - Myeloproliferative neoplasms, unclassifiable
- **Myeloid neoplasms associated with eosinophilia and abnormalities of *PDGFRA*, *PDGFRB*, or *FGFR1***
  - Myeloid neoplasms associated with *PDGFRA* rearrangement
  - Myeloid neoplasms associated with *PDGFRB* rearrangement
  - Myeloid neoplasms associated with *FGFR1* abnormalities
- **Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)**
  - Chronic myelomonocytic leukemia
  - Atypical chronic myeloid leukemia, *BCR-ABL1*-negative
  - Juvenile myelomonocytic leukemia
  - Myelodysplastic/myeloproliferative neoplasm, unclassifiable
    - *Provisional entity: refractory anemia with ring sideroblasts and thrombocytosis*
- **Myelodysplastic syndrome (MDS)**
  - Refractory cytopenia with unilineage dysplasia
    - Refractory anemia
    - Refractory neutropenia
    - Refractory thrombocytopenia
  - Refractory anemia with ring sideroblasts
  - Refractory cytopenia with multilineage dysplasia
  - Refractory anemia with excess blasts
  - Myelodysplastic syndrome with isolated del(5q)
  - Myelodysplastic syndrome, unclassifiable
  - Childhood myelodysplastic syndrome
    - *Provisional entity: refractory cytopeniaof childhood*
- **Acute myeloid leukemia and related neoplasms**
  - Acute myeloid leukemia with recurrent genetic abnormalities
    - AML with t(8;21)(q22;q22); *RUNX1-RUNX1T1*
    - AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11*
    - APL with t(15;17)(q22;q12); *PML-RARA*
    - AML with t(9;11)(p22;q23); *MLLT3-MLL*
    - AML with t(6;9)(p23;q34); *DEK-NUP214*
    - AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); *RPN1-EVI1*
    - AML (megakaryoblastic) with t(1;22)(p13;q13); *RBM15-MKL1*
    - *Provisional entity: AML with mutated* NPM1
    - *Provisional entity: AML with mutated* CEBPA
  - Acute myeloid leukemia with myelodysplasia-related changes
  - Therapy-related myeloid neoplasms
  - Acute myeloid leukemia, not otherwise specified
    - AML with minimal differentiation
    - AML without maturation
    - AML with maturation
    - Acute myelomonocytic leukemia
    - Acute monoblastic/monocytic leukemia
    - Acute erythroid leukemia
      - Pure erythroid leukemia
      - Erythroleukemia, erythroid /myeloid
    - Acute megakaryoblastic leukemia
    - Acute basophilic leukemia
    - Acute panmyelosis with myelofibrosis
  - Myeloid sarcoma
  - Myeloid proliferations related to Down syndrome
    - Transient abnormal myelopoiesis
    - Myeloid leukemia associated with Down syndrome
  - Blastic plasmacytoid dendritic cell neoplasm

1.2 Chronic Myeloid Leukemia (CML)

Chronic Myeloid leukemia (CML) is a myeloproliferative disorder characterized by excessive accumulation of apparently normal myeloid cells. It occurs with an annual incidence of 1.0 – 1.5 per 100,000 persons. CML occurs very rarely in children. In the Western word, the median age of onset is 50-60 years, which reflects the average age of the population [22]. Although symptoms at presentation may include lethargy, weight loss, unusual bleeding, sweats, anemia, and splenomegaly, in more develop countries, 50% of patients are asymptomatic and are diagnosed as a consequence of blood tests performed for unrelated reasons. More than 90% of CML patients are diagnosed when their disease is in a relatively early phase know as the chronic phase.

CML is characterized by distinct clinical phases: most patients present the disease in the chronic phase (CP), a phase in which mature granulocytes are still produced, but patients have an increased number of myeloid progenitor cells in PB. As the disease progresses, patients enter an accelerated phase (AP) followed by a blast crisis (BC), in which hematopoietic differentiation has become arrested and immature blast accumulate in the BM and spill to the circulation (Figure 2). It has been suggested that the progression of of CP to BC is a two-step process. The initial step for chronic phase is the occurrence of Philadelphia chromosome (Ph) and genetic instability caused by the BCR-ABL1 tyrosine kinase. The second step is the acquisition of additional chromosomal aberrations or mutations of transcription factors by failed DNA repair processes. The cellular and molecular mechanism in the CML progression, however, are still poorly understood.

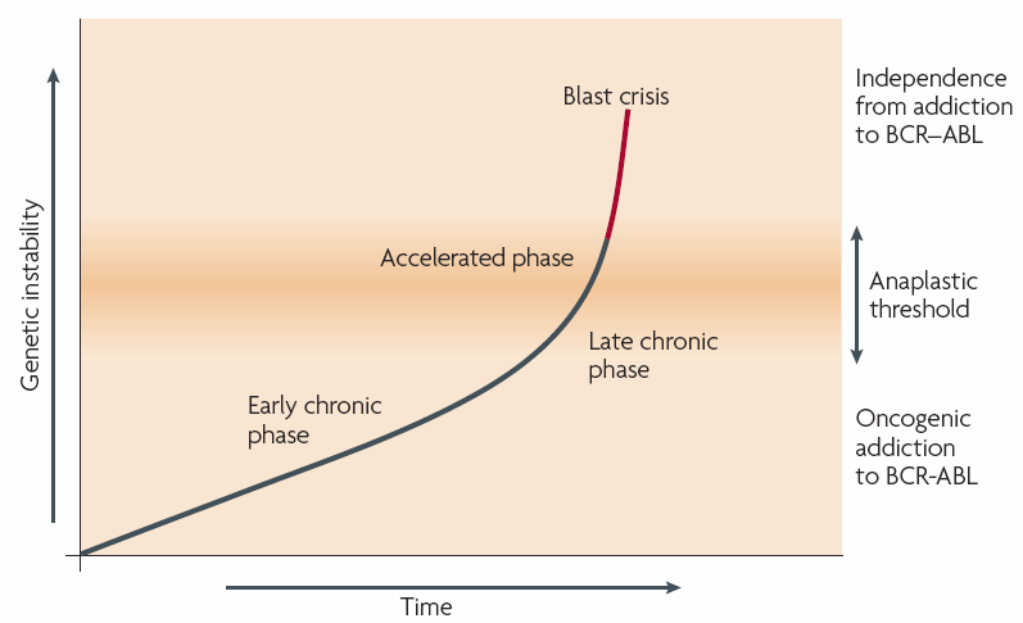

Figure 2: Natural course of chronic myeloid leukemia. In the chronic phase, the bulk of leukemic stem cells remain capable of undergoing differentiation, leading to the excessive production of mature granulocytes. In the accelerated phase, differentiation has become arrested, probably at the stage of the leukemia progenitor cell, and the 'aggressive' disease phenotype is caused by the proliferation (self-renewal) of immature blasts. Deleterious genetic events are believed to accumulate within stem and progenitor cells of the leukemic clone until there are sufficient secondary mutations to drive the transition from chronic to advanced phase disease, the blast crisis. These include: an increase in genomic instability through interference with genomic surveillance and DNA-repair proteins and a progressive telomere shortening (Nature Reviews. 2007. 7:441-53).

CP is the initial stage of CML, in which most patients are diagnosed. It usually has an insidious onset, and the main clinical findings include enlarged spleen, fatigue and weight loss. The PB shows leukocytosis (approximately 150 x 10^9/L white blood cells (WBC)), predominantely owing to neutrophils in different stages of maturation, as well as basophilia and eosinophilia. Blasts usually represent < 2% of WBC. The platelet count is normal or increased.

At the molecular level, CML-CP is characterized by the presence of the Ph chromosome and the oncogene that it encodes is present in the vast majority of myeloid cells and some lymphocytes. The Ph chromosome results from a t(9;22)(q34;q11) reciprocal translocation that juxtaposes the c-abl oncogene1 (*ABL1*) gene on chromosome 9 with the breakpoint cluster region (*BCR*) gene on chromosome 22, generating the *BCR-ABL1* fusion oncogene with a greatly enhanced ABL1 kinase activity (Figure 3). It is believed that acquisition of the BCR-ABL1 gene occurs initially in a single hematopoietic stem cell that gains aproliferative advantage and/or aberrant

differentiation capacity over its normal counterparts, giving rise to the expanded myeloid compartment.

AP is an intermediate stage of CML evolution, when the disease starts to become refractory to therapy. It is characterized by an increase in spleen size and in total WBC, blasts comprising 10-19% of the WBC, >20% circulating basophils, persistent thrombocytopenia and/or the appearance of new clonal cytogenetic abnormalities.

BC is the final stage of CML, which may or may not be preceded by an AP. Patients experience worsened performance status, and symptoms related thrombocytopenia, anemia and increased spleen enlargement. The WHO criteria for the diagnosis of BC include: blast in excess of 20% in the PB or BM; and/or extramedullary blast proliferation; and/or large foci or clusters of blasts in BM histological sections. Patients with CML-BC have a median survival of approximately 6 months.

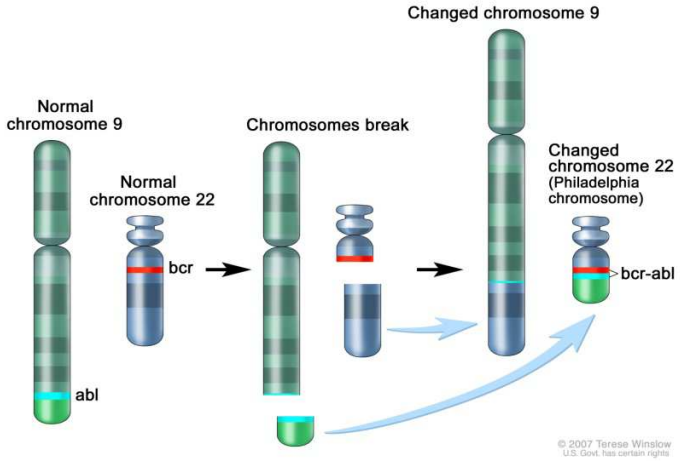

Figure 3: t(9;22) and *BCR-ABL* fusion gene. The Ph chromosome results from a t(9;22)(q34;q11) reciprocal translocation that juxtaposes the c-abl oncogene1 (*ABL1*) gene on chromosome 9 with the breakpoint cluster region (*BCR*) gene on chromosome 22, generating the *BCR-ABL1* fusion oncogene with a greatly enhanced ABL1 kinase activity (©Terese Winslow).

The natural history of the disease has damatically changed since 2000 with the emergence of tyrosine kinase inhibitors (TKI) that can block the enzymatic action of the BCR-ABL1 chimeric protein. The first of these to be developed was Imatinib®. Recent karyotype analyses show that 60%–70% of patients achieve complete disappearance of Ph-positive marrow cells and maintain exclusively Ph-negative BM cells (a state designated as a complete cytogenetic response [CCyR]) 5 years after initiating

Imatinib® treatment. The incidence of progression to a more advanced phase of leukemia in patients responding to Imatinib® is extremely low beyond the first two years [23]. However, a small number of patients fail to respond to Imatinib® (primary resistance), while others respond initially and then lose their response (secondary resistance) [24]. The reasons for Imatinib® resistance in CML-CP patients are poorly understood. Primary resistance may be related, at least in part, to the intrinsic heterogeneity of the disease (e.g., different BCR-ABL1 levels) in different patients and to the survival of variable numbers of quiescent cells from which the more mature leukemia cells are derived during CP [25]. Secondary resistance may have a wide range of causes, of which the best characterized is the acquisition of mutations in the BCR-ABL1 kinase domain (such as the T315I mutation) [26]. In the last few years, two new TKIs, Dasatinib® and Nilotinib®, have become available, both of which are more potent in vitro inhibitors of the BCR-ABL1 kinase than Imatinib®. Both of these "second-generation" TKIs are effective at inducing or restoring CCyR in 40%–50% of patients who appear to have failed primary treatment with Imatinib®. However, approximately 20% of patients presenting with CML-CP fail to respond to both Imatinib® and a subsequent second-generation TKI; their prognosis is poor because of a higher risk of disease progression. Before the advent of BCR-ABL1 TKIs, all patients with CML-CP progressed spontaneously to advanced phase CML after a median interval of approximately 5 years.

1.3 Chronic Neutrophilic Leukemia (CNL)

Chronic neutrophilic leukemia (CNL) is a rare MPN, with only ~200 patients reported to date, mostly culled from case reports and small case series. Absence of both *BCR-ABL1* and rearrangement of *PDGFRA*, *PDGFRB* or *FGFR1* are the minimal diagnostic requirements for CNL [21]. The clinical and laboratory features of CNL include hepato/splenomegaly, persistent neutrophilic leukocytosis with minimal left-shift often characterized by toxic granulation and Döhle bodies, and elevated leukocyte alkaline phosphatase (LAP) and vitamin B12 levels [21, 27-29]. According to the 2008 WHO diagnostic criteria for CNL [21], the leukocytosis is ≥ $25x10^9$/L; more than 80% of leukocytes are segmented neutrophils / band forms and <10% are immature granulocytes. Granulocytic dysplasia is not present, and there is no monocytosis, eosinophilia, or basophilia. Examination of the BM shows a myeloid hyperplasia with full maturation with <5% myeloblasts (<1% in the PB). Megakaryocytes are typically normal, but can include some small hypolobated megakaryocytes. Reticulin fibrosis is not significantly increased. Exclusionary criteria include no evidence of a reactive neutrophilia, other MPN, MDS, or overlap MDS/MPN disorder [21]. Table 2 summarizes the 2008 WHO classification for CNL.

The disease course of CNL is variable, but acceleration is typically characterized by refractory neutrophilia, worsening organomegaly, and blastic transformation. In a literature review of 40 patients meeting WHO criteria for 'true' CNL, the median survival was 23.5 months (range: 1–106) [29]. Median time to AML transformation was 21 months (range 3-94). The most frequent causes of death were intracranial hemorrhage, progressive disease/blastic transformation, and regimen-related toxicity from induction chemotherapy or transplantation [28, 29].

No standard of care exists for CNL. Therapy has primarily consisted of hydroxyurea or other oral chemotherapeutics, as well as interferon-alpha [27-34]. These agents can elicit improvement in blood counts, but exhibit no proven disease-modifying benefit. Although splenic irradiation and splenectomy may provide transient palliation of symptomatic splenomegaly, the latter has been associated with anecdotal worsening of neutrophilic leukocytosis in CNL. The limited experience with induction-type chemotherapy for blastic transformation is generally poor with death related to

resistant disease or regimen-related toxicities. Allogeneic transplantation may result in favorable long-term outcomes in selected patients, particularly when undertaken in the CP of disease [27-30, 32].

Although clonality has been demonstrated in CNL [35, 36], the majority of patients exhibit normal cytogenetics [27-29]. In CNL, trisomy 8 and del(20q) are the most common non-specific chromosomal abnormalities observed at diagnosis or at the time of progressive disease. The e19/a2 type BCR-ABL1 mRNA transcript (p230) that was initially reported as the molecular basis for some cases of CNL is instead now considered related to an uncommon 'neutrophilic variant' of CML [37]. Notwithstanding case reports of *JAK2* V617F positivity in selected cases of CNL [34, 38, 39], no other recurrent genetic mutations had been identified in these diseases until the recent discoveries of mutant *CSF3R* [38].

Table 2: WHO diagnostic criteria for CNL (adapted from Vardiman *et. al.* 2011 [40])

1	Peripheral blood leukocytosis (WBC ≥ $25x10^9$/L)
	Segmented neuthrophils and band forms are > 80% of the WBC
	Immature granulocytes (promyelocytes, myelocytes, metamyelocytes) are <10% of WBC
	Myeloblasts are <1% of WBC
2	Hypercellular BM biopsy
	Neutrophilic granulocytes increased in number and percentage
	Myeloblasts <5% of nucleated BM cells
	Neutrophilic maturation pattern normal
	Megakaryocytes normal or left shifted
3	Hepatosplenomegaly
4	No identifiable cause for physiologic neutrophilia or, if present, demostration of clonality
	No infectious or inflammatory process
	No underlying tumor
5	No Philadelphia chromosome or *BCR-ABL1* fusion gene
6	No rearrangements of *PDGFRA* , *PDGFRB* , or *FGFR1*
7	No evidence of polycythemia vera, essential thrombocythemia, or primary myelofibrosis
8	No evidence of MDS or MDS/MPN
	No granulocytic dysplasia
	No myelodysplastic changes in other myeloid lineages
	Monocytes < $1x10^9$/L

1.4 Blastic Plasmocytoid Dendritic Cell Neoplasm (BPDCN)

Blastic plasmacytoid dendritic cell neoplasm (BPDCN), previously referred as "blastic NK-cell lymphoma/leukemia" or "agranular CD4+/CD56+ hematodermic neoplasm", is a rare and aggressive hematologic disease derived from precursors of a specialized subset of dendritic cells, and hence, after some discussion, is considered a myeloid-related disease according to the 2008 WHO classification of myeloid neoplasms [21, 41-46]. These tumor cells infiltrate skin, BM, PB and lymph nodes [47-50] and mainly affect elderly patients with the maximum incidence peaking at approximately 65 years of age.

Diagnosis is based on the immunophenotype of the blast cells that are characterized by the expression of CD4, CD43, CD56, CD123, BDCA-2/CD303, TCL1, and CTLA [51]. In addition, immunohistochemical markers, such as SPIB, BDCA-4, IRF-8, BCL11A and CD2AP, have been recently reported as tools for BPDCN diagnosis [51, 52]. There are no established genetic biomarkers that assist in the clinical management of BPDCN. Preliminary results from molecular studies suggest that abnormalities of genes known to confer a poor prognosis in other hematological malignancies, such as *CDKN2A*/*CDKN2B*, *TET2* and *TP53*, are frequently seen in BPDCN [53, 54].

The overall prognosis for BPDCN is remarkably poor. Lymphoid-like chemotherapy is the preferred treatment, however, most patient relapse, resulting in a median overall survival (OS) of 12-14 months [55, 56]. Only high-dose therapy followed by allogeneic stem cell transplantation can provide durable control in this otherwise fatal condition [57, 58].

1.5 Myeloid Malignancies in the age of genomics

Since Nowell and Hungerford identified the t(9;22) translocation associated with CML [59], a wealth of data has accumulated showing that the karyotype and mutation status of certain genes provide important prognostic, and in some cases, therapeutic information for myeloid malignancies. However, standard cytogenetics has a limited resolution (3-5 megabases - Mb) to detect chromosomal abnormalities and genotyping of individual genes is expensive; therefore, as the number of mutations

identified in myeloid malignancies increases, this method may soon become impractical. During the last decade, the technology to interrogate cancer genomes has rapidly advanced. The resolution for variant detection was improved first with array-based technologies using comparative genome hybridization (CGH) and single nucleotide polymorphism (SNP) analysis, and now with whole-genome sequencing (WGS), which provides unbiased, genome-wide coverage at a single base-pair resolution. Gene-expression profiling and CGH/SNP analyses of myeloid malignancies have also provided important insights into its classification and biology [60-62].

In the past, the size and complexity of the human genome (3 billion base pairs - bp) made the sequencing of human cancer genomes impractical. Two major advances helped to overcome these obstacles. First, the generation of the draft sequence of the human genome by the Human Genome Project in 2001 provided a road map of the human genome [63]. Second, technological advances in DNA sequencing that have dramatically reduced the cost and time required to sequence genomes. Whereas the Human Genome Project took more than 10 years and several billion dollars to sequence the first human genome, current estimates are 6 weeks and $20 000 per human genome ($40 000 for paired tumor/normal genomes). Therefore, we are rapidly approaching the time when sequencing the genomes of patients with cancer will be practical in the clinical setting.

1.5.1 Next-Generation Sequencing (NGS) of Cancer

The development of massively parallel sequencing (termed next-generation or second-generation sequencing) revolutionized our ability to analyze cancer genomes. In brief, massively parallel sequencing results in the generation of millions of short (50-100 nucleotides) DNA sequences simultaneously. These sequences are then mapped back to the human reference genome to generate a picture of the cancer genome [64, 65]. For studies of cancer, it is important to sequence both the tumor and the normal tissue (oral mucosa or skin tissue, for example) from the individual. There are 3-4 million inherited sequence variants per human genome (and hundreds of copy number variants). Consequently, the majority of sequence variants identified in a cancer genome are composed by inherited polymorphisms and not by acquired mutations. Therefore, a comparison of a tumor genome with its paired normal genome is required

to efficiently identify acquired (somatic) sequence variants. Currently, there are several different ways in which NGS is being applied to study cancer genomes. The goals, advantages, and limitations of each approach are summarized below and in Table 3. Ultimately, combinations of approaches (whole-genome and transcriptome sequencing, for example) may be required to comprehensively study cancer cells.

1.5.2 Whole-Genome Sequencing (WGS)

The goal of WGS is to sequence the entire genome. Typically, this requires at least 30- to 40-fold haploid coverage of the genome (~100 Gbp of sequence) to achieve adequate diploid coverage for comprehensive mutation discovery. The advantages of WGS include: (1) the entire genome is surveyed, not just coding genes, and (2) structural variants, including deletions, amplifications, chromosomal translocations and uniparenteral disomy (UPD), are readily identified. The major limitations are cost and the complexity of the data analysis. As costs reduction and bioinformatic approaches to analyzing sequence data advance, WGS is likely to become the dominant platform for mutation detection.

1.5.3 Whole-Exome Sequencing (WES)

The goal of exome sequencing is to selectively sequence the 1%-2% of the genome containing coding genes, microRNAs, and other noncoding RNAs. In brief, genomic DNA is hybridized to beads (or arrays) containing probes designed to capture the approximately 200 000 exons in the human genome. The DNA “captured” on the beads is then subjected to next-generation sequencing. The major advantages of exome sequencing, compared with WGS, are, for instance, its reduced cost to get relatively deep sequence coverage, because only 1%-2% of the genome is analyzed. Another important advantage is the higher likelihood of a mutation that is found by WES will result in a non-properly functioning protein, making this approach an even better cost effective tool for research and clinical translation. However, exome sequencing will not detect mutations in regions outside of the exome (~ 98% of the genome) and will not detect most structural variants, such as chromosomal translocations with intronic break points.

1.5.4 Transcriptome sequencing

The goal of transcriptome sequencing is to sequence all transcribed genes [66], including both coding and noncoding RNAs. The advantages of transcriptome sequencing include: (1) quantitative information about gene-expression levels are obtained, (2) post-transcriptional changes in gene expression such as alternative splicing are detected, and (3) fusion transcripts produced by chromosomal rearrangements can be detected. Similar to sequencing the exome, transcriptome sequencing does not detect mutations in non-coding regions of the genome. It cannot detect mutations that cause the loss of one or both copies of a gene or mutations that accelerate RNA turnover (frameshift or nonsense mutations that cause "nonsense mediated decay", for example). Moreover, transcriptome sequencing is biased toward abundantly expressed transcripts and therefore the sequencing coverage of genes expressed at lower levels can be below or absent.

1.5.5 Other sequencing applications

There are emerging NGS applications to characterize genome-wide epigenetic modifications in cancer cells. This is particularly relevant to the study of AML, because genes regulating DNA methylation are frequently mutated in AML, including *DNMT3A*. NGS techniques have been developed to map genome-wide DNA methylation at a single-base resolution [67] and to assess chromatin structure by mapping the location of post-translationally modified (methylated and acetylated, for example) histones on the genome [68].

Table 3: Summary of next-generation sequencing assays

Sequencing target	Approach	Comment
Whole genome	Unselected genomic DNA is fragmented and sequenced	Comprehensive but expensive; detects point mutations, indels, translocation, deletions, and amplifications
Exome	Exonic DNA is enriched by capture and then sequenced	Only identifies mutations in exons, misses many structural variants
Transcriptome	RNA is reversed transcribed to generate cDNA and then sequenced	RNA expression and splicing are assessed; only identifies mutations in expressed genes
Methylome	gDNA is treated with bisulfite to mark methylated cytosines before sequencing	Assay being optimized; provides genome-wide information about gDNA methylation
Histone-associated genome	ChIP using antibodies against modified histones followed by sequencing	Used to assess chromatin modifications, in particular genome-wide occupancy by modified histones

1.5.6 Current limitations of whole-genome sequencing

Although clearly a valuable and promising approach to analyzing myeloid malignancies (and other cancers), there are some technical limitations of whole-genome sequencing. The ability of next-generation sequencing to detect sequence variants is dependent on the "read-depth," the number of unique times a single nucleotide is sampled. Whole-genome sequencing typically aims to cover each nucleotide with an average of 30-40 separate reads. However, at this coverage, nearly 50% of sequence variants identified are false positives due to mapping errors, polymerase errors, and the low frequency of somatic variants within the sample. Therefore, secondary validation of variants remains a necessary part of all sequencing projects. The detection of small insertions and deletions (indels) using current algorithms is problematic. In addition, mapping sequence reads to the approximately 50% of the human genome that contains repetitive sequences is imperfect, raising the possibility that mutations are being missed. The incident AML case highlighted these technical issues [69]; a single base-pair deletion in *DNMT3A* was missed during short-read analysis (30- to 40-bp reads, which are difficult to align to the reference genome if they contain an indel), but was later identified when the genome was re-sequenced with paired-end reads (pairs of 50- to 75-bp reads, which align more robustly). These problems will be improved as bioinformatics can draw on additional sequenced human genomes and as decreased costs permit deeper sequencing, but efficiently mapping sequences to repetitive regions of the genome will likely require technical improvements that permit longer sequence read lengths.

There are inherent limitations of whole-genome sequencing that need to be considered when analyzing cancers. First, each leukemia genome carries approximately 500-1000 somatic mutations [70]. It is likely that the majority of these mutations do not contribute to leukemic transformation. Identifying the "driver'" mutations in each leukemia genome is currently challenging. Second, whole-genome sequencing does not provide information about epigenetic modifications and alterations in gene expression. Therefore, to fully characterize acquired genomic changes in a patient, assays to measure RNA expression (transcriptome sequencing,

for example) and epigenetic changes (e.g., bisulfate sequencing to define the methylome) in combination with whole-genome sequencing are still required.

Genome sequencing remains expensive, and the requisite infrastructure, expertise, and time to complete sequence analysis are significant barriers to the routine use of WGS in the clinical setting. However, as the number of sequenced genomes increases, the technical and bioinformatic infrastructure required for clinical analysis will become more accessible. Current cost estimates at the Washington University Genome Institute for WGS of a paired normal/leukemia sample is $40 000. This is a fully loaded cost that includes sequencing, bioinformatic analysis, and validation. The current time to sequence and analyze a paired tumor and normal genome is 6 weeks. The cost of WGS needs to be considered in the context of the increasingly complex diagnostic evaluation for leukemia. A common molecular evaluation of AML consists of cytogenetics; FISH for t(15;17), t(8;21), inv(16), and t(9;22); and molecular testing for *FLT3*, *KIT*, *JAK2* and *NPM1*. The list price of this "standard of care" molecular evaluation at a laboratory is approximately $4500 and does not include the cost of molecular testing for other recurring mutations (e.g., *NRAS, IDH1, IDH2, TET2*, *DNMT3A*, *RUNX1, ASXL1*, and *CEBPA*) that are likely to become routine. As the cost of sequencing continues to decrease and the number of molecular assays continues to increase, we will rapidly reach the inflection point where NGS will be the most cost-effective diagnostic tool for leukemia patients.

Finally, there are important practical and ethical considerations concerning the clinical application of WGS. Based on the cancer genomes analyzed to date, it is clear that a tremendous number of genetic variants will be identified in each genome, of which the majority will be inherited variants. Some of these will be in genes known to be involved in cancer susceptibility, and some will be associated with unsuspected genetic diseases (e.g., hereditary hemochromatosis). Best practices to communicate this information to patients and their families are being considered and developed by several groups of investigators.

1.5.7 Potential clinical applications (current and future)

NGS provides a method to identify all genetic changes in a cancer genome, including single-nucleotide mutations, deletions, amplifications, chromosomal

translocations, and UPD, at a single-base resolution. There is several ways that this unprecedented ability to characterize cancer genomes may affect the clinical management of patients with leukemia.

- Classification

NGS may lead to a better classification of leukemias. There is evidence that a molecular classification of leukemia based on karyotype and specific gene mutations is superior to a morphology-based classification. As additional mutations are discovered in leukemias, it is likely that the molecular classification of this disease will be further defined. Having access to all of the genetic changes in a genome through WGS would facilitate this process.

- Optimizing therapy

A detailed knowledge of mutations present in leukemia genome may help to optimize the therapeutic plan. For example, mutations in tyrosine kinase genes may suggest the use of specific small molecule inhibitors. In addition, disruption of specific pathways in the leukemia genome may be exploited therapeutically. For example, mutations in *BRCA1* or *BRCA2*, by inhibiting DNA repair through homologous recombination, may render cancer cells more sensitive to inhibitors of protein poly(ADP)-ribose polymerase [71]. Finally, the discovery of new mutations in leukemia may lead to the development of novel targeted therapeutics.

- Pharmacogenomics

Pharmacogenomics is the study of how an individual's genes influence his or her response to drugs. Genetic variants can influence the metabolism, toxicity, and therapeutic efficacy of a drug. By sequencing the "normal" and leukemia genomes in patients with cancer, genetic variants known to affect the efficacy and toxicity of chemotherapeutic agents will be identified and this information may affect the choice and dose of chemotherapy used. For example, germline polymorphisms in the UDP glucuronosyltransferase1 family (*UGT1A1*) increase the toxic effects of irinotecan [72], and germline polymorphisms in thiopurine S-methyltransferase (*TMPT*) increase sensitivity to mercaptopurine through decreased metabolism [73]. As more genomes

are sequenced, the list of genetic variants influencing response to chemotherapy likely will expand.

- Cancer susceptibility

As discussed above, WGS of the "normal" genome in patients with cancer may identify genetic variants contributing to cancer susceptibility. Not only will this information be important for genetic counseling, it may affect the patients treatment and prognosis. WGS will provide genotyping data for the more than 100 genes already implicated in cancer susceptibility. Causal associations between germline variants and AML have been firmly established for only a few genes, such as *TP53* [74] and *RUNX1* [75]. However, it is likely that the list of genetic variants contributing to leukemia susceptibility will increase in the near future as additional genomes are sequenced.

1.5.8 Identification of novel somatic mutations in myeloid malignancies by NGS

Traditional approaches to mutation discovery in cancer rely on sequencing of a pre-selected group of genes. Typically, genes or gene family members already implicated in cancer are sequenced. This directed sequencing approach (in effect "looking under the lamppost"), while yielding important insights into cancer biology is inherently limited. In contrast, genomic approaches study the entire cancer genome and thus provide a nonbiased way to look for mutations. Moreover, by simultaneously evaluating all genes, WGS is an ideal platform to assess cooperativity between gene mutations.

In a study published in 2008, Ley et al reported the sequence of the first primary cancer genome, a case of FAB M1-AML [69]. Since then, the sequence of 3 additional AML genomes have been reported: an additional case of M1-AML (normal karyotype) [69, 70, 76] and 1 case each of M3-AML [77] and therapy-related AML (complex karyotype) [78]. In addition, exome sequencing was reported in 14 cases of M5-AML [79].

These studies have begun to define several important global features of AML. First, there are only a relatively small number of somatically acquired single nucleotide variants (SNVs or "point mutations") in AML cells. Within coding regions, the number

of somatic SNVs ranged from 10-26 in the 4 sequenced AML genomes [69, 70, 76-78]. This suggests that AML does not result from genomic instability, but rather may be associated with a small number of oncogenic “driver” mutations. Second, most acquired AML associated SNVs are private: of the 59 reported SNVs in the coding regions of these 4 genomes, no gene was mutated in 2 or more patients. This suggests that many of these mutations are likely to be personal “passenger” mutations, which are random and do not contribute to AML pathogenesis.

Despite the lack of recurring mutations within these 4 genomes, these studies identified 2 novel AML-associated genes that were found to be frequently mutated in AML and contribute to leukemogenesis. Mardis et al identified mutations in *IDH1* in the second AML genome published and found an additional 16 patients carrying *IDH1* mutations [76]. Mutations in *IDH1* and *IDH2* have since been identified in AML patients in several studies, occur at a frequency of 7%-16%, and appear to be mutually exclusive of one another and of acquired *TET2* mutations [80, 81]. A recent study found that mutations in *IDH1* or *IDH2* disrupted *TET2* function and led to a hypermethylation phenotype with impaired hematopoietic differentiation [82], although a second study found impaired methylation associated with *TET2* mutations [83]. Efforts are ongoing to develop targeted therapeutics that might abrogate these effects.

Acquired mutations in *DNMT3A*, a de novo DNA methyltransferase, were identified by NGS of AML samples by 3 independent groups. Yamashita et al used targeted-exome sequencing to identify acquired mutations in *DNMT3A* at amino acid R882 in 3 of 71 (4.2%) AML cases [84]. Ley et al identified a single base-pair deletion in *DNMT3A* by WGS and identified acquired mutations (including deletions, frameshifts, nonsense, and missense mutations) in an additional 62 of 281 AML patients (22.1%) [70]. They showed that *DNMT3A* mutations were more common in AML with normal cytogenetics (33.7%) or M5-AML cytomorphology (57.1%) and were associated with an adverse outcome. Finally, Yan et al used WES to identify missense mutations at R882 and at 3 other amino acid positions in 23 of 112 (20.5%) patients with AML-M5 [79]. It is currently unclear whether some of these mutations (e.g., the highly recurrent R882 mutations) might be associated with a gain of function.

In 2013, the Cancer Genome Atlas Research Network analyzed the genomes of 200 clinically annotated adult cases of de novo AML, using either WGS (50 cases) or WES (150 cases), along with RNA and microRNA sequencing and DNA-methylation analysis. AML genomes have fewer mutations than most other adult cancers, with an average of only 13 mutations found in genes. Of these, averages of 5 are in genes that are recurrently mutated in AML. A total of 23 genes were significantly mutated, and another 237 were mutated in two or more samples. Nearly all samples had at least 1 nonsynonymous mutation in one of nine categories of genes that are almost certainly relevant for pathogenesis, including transcription-factor fusions (18% of cases), the gene encoding nucleophosmin (NPM1) (27%), tumor-suppressor genes (16%), DNA-methylation-related genes (44%), signaling genes (59%), chromatin-modifying genes (30%), myeloid transcription-factor genes (22%), cohesin-complex genes (13%), and spliceosome-complex genes (14%). Patterns of cooperation and mutual exclusivity suggested strong biologic relationships among several of the genes and categories. Their study identified at least one potential driver mutation in nearly all AML samples and found that a complex interplay of genetic events contributes to AML pathogenesis in individual patients [85].

Regarding others myeloid disorders, in 2011, Bejar *et. al.* [86] used a combination of genomic approaches, including NGS and mass spectrometry-based genotyping, to identify mutations in samples of BM aspirate from 439 patients with MDS. They identified somatic mutations in 18 genes, including two, *ETV6* and *GNAS* that have not been reported to be mutated in patients with MDS. A total of 51% of all patients had at least one point mutation, including 52% of the patients with normal cytogenetics. Mutations in *RUNX1*, *TP53*, and *NRAS* were most strongly associated with severe thrombocytopenia and an increased proportion of BM blasts. In a multivariable Cox regression model, the presence of mutations in five genes retained independent prognostic significance: *TP53*, *EZH2*, *ETV6*, *RUNX1* and *ASXL1*.

CHAPTER 2: Aims

The Human Genome Project and the advances in sequencing technologies have revolutionized our ability to characterize cancers at the molecular level. Using next-generation sequencing (NGS), it is now possible to interrogate cancers genome wide at a single-base-pair resolution. Application of NGS to myeloid leukemia has already yielded important discoveries, including the identification of common gene mutations (eg, *IDH1* and *DNTM3A*). However, these efforts have so far excluded rare myeloid disorders. Taking all this into account, we proposed the next aims:

1. Evaluate if a systematic genomic strategy based on whole-genome sequencing, RNA sequencing and target re-sequencing approaches could provide a deeper insights into the genetic pathogenesis of rare myeloid disorders.
2. If so, to evaluate how these technologies can be used to improve the clinical management of the affected patients. In detail:

   2.1 In a rare case of blastic transformation of chronic myeloid leukemia

   - Identify the genetic changes associated with this transformation by whole-exome sequencing
   - Study additional alterations in malignant cells (Ph+), related with no response to therapy, and their frequency in a validation cohort

   2.2 In a case of chronic neutrophilic leukemia (CNL)

   - Identify the genetic changes associated with CNL pathogenesis by whole-exome sequencing
   - Identify the CNL transcriptome profiling and determine the functional consequences of *LUC7L2* and *U2AF1* mutations in splicing by RNA sequencing

   2.3 In a series of patients with blastic plasmacytoid dendritic cell neoplasm (BPDCN)

   - Identify the genetic changes associated with BPDCN pathogenesis by whole-exome sequencing
   - Design a comprehensive and cost-effective target re-sequencing approach based on the WES results and study the frequency of mutations in a validation cohort
   - Evaluate the impact of mutation in the patients outcome

CHAPTER 3: Material and Methods

3.1 Patients samples

All patients samples included in this work fulfilled the 2008 WHO classification criteria for myeloid neoplasm [21]. A proper inform consent was obtained from each patient and the study was approved by the ethical research and animal care committee from Institute of Health Carlos III (CEi PI 32_2009).

Data collection comprised the following clinical and analytical parameters: birth date, gender, hospital of origin, diagnosis date, histological subtype, extra medullary affection, stage, hemoglobin, white blood cell count, lymphopenia (lymphocytes <600), leukocytosis (leucocytes > 15000), IPS, treatment, type of response (CR: complete response; PR: partial response; P: progression) and data regarding last follow up (dead, alive with or without disease).

The sample came from 10 clinical and basic research groups in Spain: Hospital La Paz (Madrid), Complejo Hospitalario Ntra. Sra. de Candelaria (Sta. Cruz de Tenerife), Hospital General Universitario Gregorio Marañón (Madrid), Hospital Universitario La Fe (Valencia), Hospital Universitario Marqués de Valdecilla (Santander), Hospital Universitario de Salamanca (Salamanca), Hospital Universitario 12 de octubre (Madrid), Universidad de Navarra (Pamplona), Hospital del Mar (Barcelona) and Hospital Niño Jesus (Madrid).

Different patient series were used in this work and considered as follows:

3.1.1 Chronic Myeloid Leukemia (CML) project

Samples were obtained from normal tissue (oral mucosa) and tumor bone marrow (BM) blasts at three different time-points (CP, complete hematological remission-CHR and BC) of a index patient with CML. These samples were used as a discorevy cohort to identify the genetic changes associated with transformation to the aggressive phenotype of blastic crisis by whole-exome sequencing (WES). The candidate mutations were analyzed in 26 additional CML samples: 13 at CP-CML paired with 13 follow up/ BC- CML (extension cohort). A summary of the clinical characteristics of CML patients can be found in Table 9 and in the chapter 4.1.1 of this thesis.

3.1.2 Chronic Neutrophilic Leukemia (CNL) project

Samples were obtained from normal tissue and BM of an index patient with CNL before treatment. The detailed clinical characteristic of this patient is described in chapter 4.2.1 of this thesis. After cytogenetic and molecular studies for diagnosis purpose, these samples were used for WES and RNA sequencing studies.

3.1.3 Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN) project

Samples for WES were obtained from normal tissue and tumor BM from three patients with BPDCN (discovery cohort) at diagnosis. Thirty-nine additional tumors were further included for the targeted re-sequencing analysis: 11 tumor samples came from BM cells, and 28 samples from formalin fixed paraffin embedded tissues (FFPE) (extension cohort). Clinical characteristics of the patients are summarized in Table 4.

3.2 Molecular cytogenetics

Cytogenetic analysis was performed on metaphases spreads obtained from the BM sample, as previously described [87]. The chromosomes were stained using the G-banding technique and 20 metaphases were analyzed using a conventional microscope and the IKAROS-software (Metasystems). Clonal chromosomal abnormalities were described according to the International System of Human Cytogenetic Nomenclature (ISCN, 2013).

Metaphase spreads obtained from the patient were hybridized overnight at 37°C with commercial diagnosis probes. Following post-hybridization washes, the chromosomes were counterstained with DAPI in antifade solution (Vector Labs). Cell images were captured using a CCD camera (Photometrics SenSys camera) connected to a computer running the Chromofluor image analysis system (Cytovision, Applied Imaging Ltd, Newcastle, UK).

Table 4: Summary of the clinical characteristics of BPDCN patients

					Tumour Phenotype							Sites of involvement at Diagnosis					
ID	**Sample**	**Site**	**Age (years)**	**Sex**	**CD4**	**CD56**	**CD123**	**TCL1**	**CD19**	**CD3**	**CD33**	**Skin**	**Bone Marrow**	**Peripheral Blood**	**Lymph node**	**Other sites**	**Treatment**
1*	D	BM	66	M	+	+	+	-	NA	NA	+	n	y	n	n	n	Dasatinib
2*	D	BM	68	M	+	+	+	NA	NA	NA	+	y	y	n	n	n	AML-like
3*	D	BM	31	F	+	NA	NA	NA	NA	NA	+	y	y	n	n	n	AML-like
4	D	BM	14	M	+	+	+	NA	-	-	+	y	y	y	y	y	ALL/Lymphoma-like
5	D	BM	56	M	+	+	+	NA	-	-	+	n	y	y	n	n	AML-like
6	D	BM	71	M	+	+	NA	NA	-	-	+	y	y	y	n	n	ALL/Lymphoma-like
7	D	BM	57	M	+	+	+	NA	-	-	-	n	y	y	y	y	ALL/Lymphoma-like
8	D	BM	64	M	+	-	+	NA	-	-	-	y	y	y	y	y	ALL/Lymphoma-like
9	D	BM	63	F	+	-	+	NA	-	-	-	y	y	y	n	n	None
10	D	BM	28	F	+	+	+	NA	-	-	-	y	y	n	n	y	ALL/Lymphoma-like
11	D	BM	39	F	NA	-	+	NA	-	NA	+	n	y	y	n	n	AML-like
12	D	BM	41	F	+	-	+	NA	-	-	-	n	y	y	y	n	ALL/Lymphoma-like
13	D	BM	35	M	+	-	+	NA	-	-	-	y	y	y	n	n	ALL/Lymphoma-like
14	D	SP	74	M	+	+	+	NA	-	-	-	y	y	y	y	y	None
15	D	BM	69	M	+	+	+	+	NA	-	NA	y	y	NA	n	n	ALL/Lymphoma-like
16	D	NA	NA	NA	+	+	+	+	NA	-	NA	NA	y	NA	NA	NA	NA
17	D	S	62	M	+	+	+	+	NA	-	NA	y	y	NA	y	y	AML-like
18	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
19	D	S	65	M	+	+	+	+	NA	-	NA	y	y	NA	y	y	ALL/Lymphoma-like
20	D	BM	29	F	+	+	+	+	NA	-	NA	NA	y	NA	n	y	ALL/Lymphoma-like
21	D	S	12	F	+	+	+	+	NA	-	NA	y	n	NA	y	n	ALL/Lymphoma-like
22	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
23	D	S	76	M	+	+	+	+	NA	-	NA	y	n	NA	y	n	None
24	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
25	D	S	59	M	+	+	+	+	NA	-	NA	y	y	NA	y	NA	ALL/Lymphoma-like
26	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
27	D	Ly	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
28	D	S	43	F	+	+	+	+	NA	-	NA	y	n	NA	y	n	ALL/Lymphoma-like
29	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
30	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
31	D	S	79	M	+	+	+	+	NA	-	NA	y	y	NA	n	n	ALL/Lymphoma-like/ AML-like
32	D	S	45	F	+	+	+	+	NA	-	NA	NA	NA	NA	y	y	ALL/Lymphoma-like
33	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
34	D	S	79	M	+	+	+	+	NA	-	NA	y	n	NA	n	n	ALL/Lymphoma-like
35	D	S	61	M	+	+	+	+	NA	-	NA	y	y	NA	y	y	ALL/Lymphoma-like
36	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	y	NA	y	NA	NA
37	D	Ly	7	M	+	+	+	+	NA	-	NA	y	n	NA	y	n	ALL/Lymphoma-like
38	D	S	80	M	+	+	+	+	NA	-	NA	y	y	NA	n	NA	ALL/Lymphoma-like
39	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
40	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
41	D	S	NA	NA	+	+	+	+	NA	-	NA	NA	NA	NA	NA	NA	NA
42	D	S	13	F	+	+	+	+	NA	-	NA	y	n	NA	n	n	ALL/Lymphoma-like

*Whole-exome sequencing; BM: bone marrow; S: skin; Ly: lymph node; PB: peripheral blood; D: diagnosis; CR: complete remission; R: relapse; BPDCN: blastic plasmocitoid dendritic cell neoplasm; M: male; F: female; NA: not available; y: yes; n: no; NoR: no response

3.3 Whole-Exome Sequencing (WES)

DNA was extracted by a phenol–chloroform standard method and the quality and quantity of purified DNA was assessed by fluorometry (Qubit, Invitrogen) and gel electrophoresis.

The preparation of shotgun libraries from the leukemic and non-leukemic genomic DNA followed the "SureSelect Human All Exon Kit" protocol (Agilent Technologies, Santa Clara, CA), covering 50 Mb of coding exons (approximately 1.60% of the genome). Briefly, 1–3 µg of high molecular weight genomic DNA from each sample was fragmented by acoustic shearing on a Covaris S2 instrument. Fractions of 150–300 bp were ligated to Illumina adapters and PCR-amplified for 6 cycles. Exon Enrichment: 300 ng of whole library were hybridized to SureSelect Human All Exon Capture kit for 24 h at 65ºC. Biotinylated hybrids were captured, and the enriched libraries were completed with 12 cycles of PCR. The resulting purified DNA library was applied to an Illumina flow cell for cluster generation and sequenced using the Illumina Genome Analyzer IIx for 76 bases in a paired-end format by following the manufacturer protocols.

3.4 WES bioinformatic analysis

The bioinformatics analysis began from the sequencing raw data (.fastq files) which generated from the Illumina pipeline. Sequenced reads were quality-checked with FastQC [88]. First, the adapter sequence in the raw data was removed, and low quality reads which have too many Ns and low base quality bases were discarded. This step produced the "clean data". Second, a combination of *Burrows-Wheeler Aligner* (*BWA*) and *Blat-like Fast Accurate Search Tool* (*BFAST*) were used to do the short-reads alignment. For the mapping we use the human genome build37 (hg19) as the reference genome. *BWA+BFAST* provide alignments in BAM format. Base qualities were recalibrated using the Genome Analysis Tool Kit (*GATK*) v1.03540 (http://www.broadinstitute.org/gsa/wiki/ index.php/The_Genome_Analysis_Toolkit) and duplicate fragments marked using *Picard Tools* v1.17 (http://picard.sourceforge.net/). After these processes, the final BAM files used to do the variant calling got ready. The BAM files were also used to visualize aligned reads

using *IGV* (Integrative Genomics Viewer) program (http://www.broadinstitute.org/igv/). We used *GATK library* to detect Single Nucleotide Polymorphisms (SNPs) and to annotate the confident variant results. The resulting call sets in the target regions (50Mb) with a sequence read depth of ≥ 10x were reported. The final variants can feed to the downstream advanced analysis pipeline. The SNP Effect Predictor tool was used to quickly and accurately predict the effects of variants on Ensembl-annotated transcripts (http://www.ensembl.org/Homo_sapiens/UserData/UploadVariations) [89].

We selected the single nucleotide variations (SNVs) and small insertions and deletions (indels) filtered them by discarding (i) variants present in the sequenced matched normal cells of the patient or in the dbSNP132 database and (ii) variants in intergenic or intronic regions. Next, we selected only those variants within coding regions not predicted to produce synonymous amino acid changes and presenting sufficient depth and quality. After that, we checked candidate variants against the 1000 Genomes Project data. The percentage of reads supporting the mutation of the total number of reads at a given position was taken as >15% in the tumor DNA, and 0.5% in normal DNA [90]. Data has been deposited in the Sequence Read Archive (SRA) database. In order to estimate the functional consequences of the missense mutations, we used two different algorithms: SIFT and PolyPhen-2.

Figure 4 represents a workflow scheme form WES procedure and bioinformatic analysis.

3.5 Ingenuity Pathway Analysis (IPA)

Genes found to be mutated among the cases were subjected to Ingenuity Pathways Analysis (IPA) (Ingenuity Systems) and used as a starting point for building biological networks (http://www.ingenuity.com/products/ipa). IPA uses the proteins from the highest-scoring network to extract a connectivity pathway that relates candidate proteins to each other based on their interactions. The involved Function & Disease and Canonical Pathways were shown to be significantly associated with these candidates.

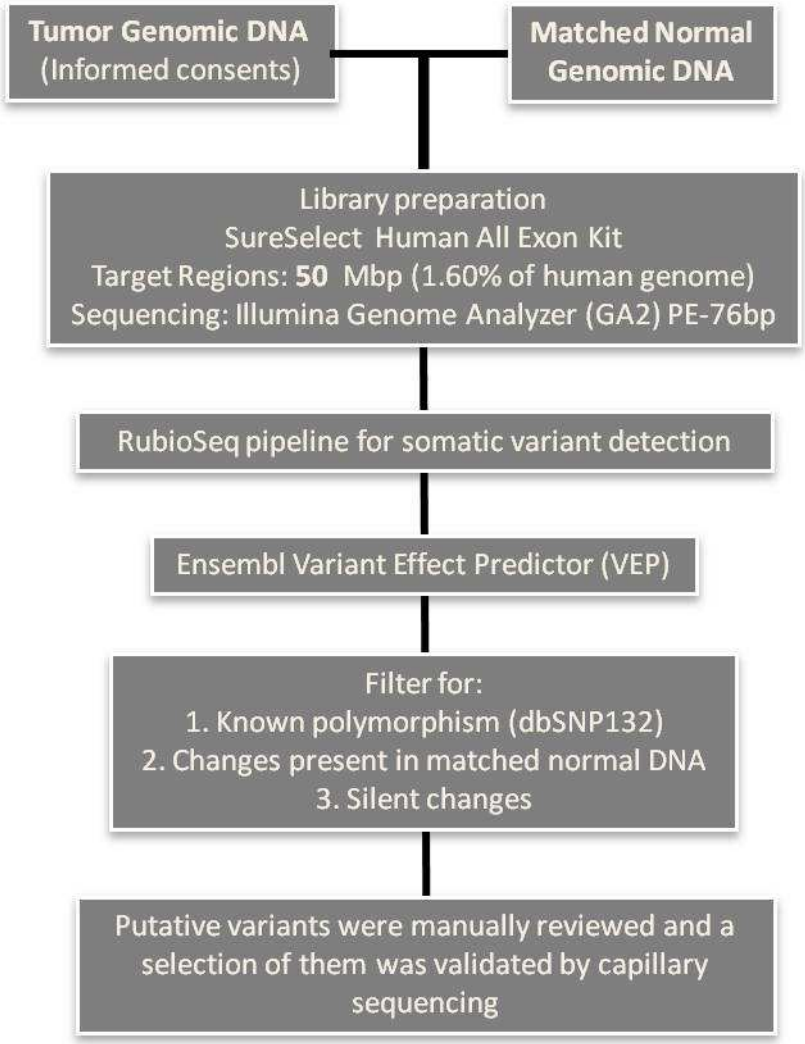

Figure 4: Workflow scheme for exon enrichment and WES analysis. After alignment of the sequencing reads from DNA samples of tumor and normal cells to the human reference genome, a series of filters were applied to discard reads that were not usable for the downstream purpose of somatic mutation discovery. Sequence variants fulfilling the additional criteria were subjected to Sanger sequencing validation.

3.6 Targeted Re-Sequencing

We design a custom panel of 38 genes (Table 5) using SureDesign Tool (https://earray.chem.agilent.com/suredesign/), which are prone to harbor mutations involved in leukemogenesis (n=25) and have been identified in our WES previous analysis (n=13). This panel is composed of genes from nine categories: DNA methylation (5 genes), chromatin remodeling (4 genes), transcription factors (10 genes), splicing (5 genes), protein kinase (6 genes), ubiquitination (4 genes) and three additional singular genes (9 categories) .Ion semiconductor sequencing on the Ion Torrent Personal Genome Machine (PGM) was performed using the HaloPlex PCR target enrichment system (Halo Genomics, Uppsala, Sweden). In short, patient DNAs were digested by a cocktail of restriction enzymes and hybridized to specific probes incorporating Ion Torrent specific sequence motifs. Hybridized molecules were captured with magnetic beads, subsequently amplified using indexed primers (Ion One Touch Template Kit), and finally sequenced with the Ion Sequencing Kit v2.0 on an Ion 316 chip. Data were analyzed with Variant Caller v2.2.3-31149 (all Life Technologies, Carlsbad, CA).

3.7 Sanger sequencing validation and screening

We used polymerase chain reaction (PCR) amplification and direct DNA sequencing to validate candidate variants. Because of the sensitivity of Sanger, sequence variants that were reported in less than 20% of the reads could not be included in this validation phase. The purified products were subsequently sequenced using the automatic sequencer ABI 3730xl (PE Applied Biosystems, Foster City, CA). Data were analyzed with BioEdit Sequence Alignment Editor 7.0.5.3. Sequences of primers used are listed on chapter 4 of this thesis according to each project.

PCR and direct DNA Sanger sequencing of *IKZF3*, *ASXL1* and *TP53* were performed in 13 paired diagnosis and follow-up/BC samples. Primers spanning all coding exons of *IKZF3* (ENST00000346872), exon 13 of *ASXL1* (ENST00000375687) and all coding exons of *TP53* (ENST00000269305) are reported in Tables 6.

Additionally, *FLT3-ITD* detection was carried out by conventional PCR in all samples as previously reported [91].

3.8 RNA Sequencing (RNA seq)

Total RNA was extracted from BM leukemia cells and from CD34+ cells (normal control) by TRIzol (Life Technologies, Carlsbad, CA), following the manufacturers protocol; 250-1000 ng of total RNA were used for the synthesis of cDNA libraries with TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA) according to the manufacturers recommendations. Sequencing by synthesis was performed on HiScanSQ sequencer (Illumina) at 75bp in paired end.

3.9 RNA seq bioinformatics analysis

Sequenced reads were quality-checked with FastQC [88]. The 75-nt paired-end reads were aligned to the human genome (GRCh37/hg19) with TopHat-2.0.4 [92] (using Bowtie 0.12.7 [93] and Samtools 0.1.16 [94]), allowing two mismatches and five multihits. Fusion transcripts were obtained with TopHat-Fusion [95].

Transcripts differential expression between samples was calculated with Cuffdiff (Cufflinks 1.3.0) [92], using the human genome annotation data set Homo_sapiens.GRCh37.65 from Ensembl.

3.10 Real-time PCR validation

LUC7L2 expression was validate by quantitative real-time RT-PCR starting from total RNA from the patient sample, which was reverse-transcribed using random hexamer primers with the TaqMan® Gold RT-PCR Kit (Hs00255388_m1 - Applied Biosystems, Foster City, CA). The Calibrated Normalized Relative Quantity, taking into account target- and run-specific amplification efficiencies, was calculated using endogenous GAPDH expression with the Qbase software application. PIM3-SCO2 fusion transcripts were amplified (PIM3-2F: CGCGACATTAAGGACGAAAA / SCO2-2R: GCCCTGCCTTGACAAAAG) and sequenced with BigDye terminator v3.1 Cycle Sequencing kit (PE Applied Biosystems, Foster City, CA) on an Applied Biosystems 310 analyzer.

Table 5: Genes included in the target next generation sequencing

N=	Gene	Coverage	Source
1	ASXL1	100.0%	CCDS13201.1
2	ATRX	100.0%	CCDS14435.1, CCDS14434.1
3	CBL	100.0%	CCDS8418.1
4	CBLB	100.0%	CCDS2948.1
5	CBLC	100.0%	CCDS12643.1, CCDS46109.1
6	CEBPA	100.0%	CCDS54243.1
7	DNMT3A	100.0%	CCDS1718.2, CCDS33157.1, CCDS46232.1
8	ETV6	100.0%	CCDS8643.1
9	EZH2	96.4%	CCDS5892.1, CCDS5891.1
10	FLT3	100.0%	CCDS31953.1
11	HOXA1	100.0%	CCDS5401.1, CCDS5402.2
12	HOXB9	100.0%	CCDS11534.1
13	IDH1	100.0%	CCDS2381.1
14	IDH2	100.0%	CCDS10359.1
15	IKZF1	100.0%	NM_001220773, NM_001220767
16	IKZF2	100.0%	CCDS2395.1, CCDS46507.1
17	IKZF3	100.0%	CCDS11346.1, CCDS11351.1, CCDS11349.1
18	JAK2	100.0%	CCDS6457.1
19	KDM6A	100.0%	CCDS14265.1
20	KIT	100.0%	CCDS3496.1, CCDS47058.1
21	KRAS	100.0%	CCDS8702.1, CCDS8703.1
22	LUC7L2	100.0%	CCDS43656.1
23	MPL	100.0%	CCDS483.1
24	NBL1	100.0%	CCDS41278.1, CCDS196.1
25	NPM1	100.0%	CCDS4376.1, CCDS43399.1, CCDS4377.1
26	NRAS	100.0%	CCDS877.1
27	PDGFRA	100.0%	CCDS3495.1
28	RUNX1	99.9%	NM_001001890, NM_001754, NM_001122607
29	SF3B1	99.9%	CCDS33356.1, CCDS46479.1
30	SRSF2	100.0%	CCDS11749.1
31	TET1	100.0%	CCDS7281.1
32	TET2	100.0%	CCDS47120.1, CCDS3666.1
33	TP53	100.0%	CCDS11118.1, CCDS45605.1, CCDS45606.1
34	U2AF1	100.0%	CCDS33574.1, CCDS13694.1, CCDS42948.1
35	UBE2G2	98.2%	CCDS33586.1, CCDS13714.1
36	WT1	100.0%	CCDS44562.1, CCDS7878.2, CCDS44561.1
37	ZEB2	100.0%	CCDS54403.1, CCDS2186.1
38	ZRSR2	97.5%	CCDS14172.1

Number of target region: 469
Total target region size: 86266 bp
Coverage: 99,9% (86137 bp)

Table 6: Primers for *ASXL1*, *IKZF3* and *TP53* screening

Primer name	Sequence (5'-3')
ASXL1-13F	GGACCCTCGCAGACATTAAA
ASXL1-13R	AGCTCTGGACATGGCAGTTC
ASXL1-13aF	GGTCAGATCACCCAGTCAGTT
ASXL1-13aR	CACCACCATCACCACTGCT
ASXL1-13bF	CAACTACTGCCGCCTTATCC
ASXL1-13bR	GTCGGTGAGGATTCAGGTGT
ASXL1-13cF	TGAAGGATCCTGTAAATGTGACC
ASXL1-13cR	ATTGCTGTCACTGCCTCCTC
ASXL1-13dF	TCACTCTGGACTGTGCCATC
ASXL1-13dR	GCAGCAACTGCATCACAAGT
ASXL1-13eF	CTCAGTGGAGGCCACTAACC
ASXL1-13eR	CATAGCACGGACTTCCTTCTG
ASXL1-13fF	CAGAAGGAAGTCCGTGCTATG
ASXL1-13fR	GGGACTATGCCCAGTAGCTTT
ASXL1-13gF	GGGTCCTCTTAAGGCAAATG
ASXL1-13gR	GTTTCCCATGGCCATAATTT
IKZF3-2F	GCTGTTAAGTTTTCCGAAATGG
IKZF3-2R	AAGCACATTCCTCTCTCTTTGG
IKZF3-3F	TGTGTGAAGCTGAAAGTTAAATGG
IKZF3-3R	CTACAATTGCAAGTTTTCCTTGG
IKZF3-4F	TCGTTGTTCATTTCTTGCATTT
IKZF3-4R	TTCTCACGTGGCTGCATTAG
IKZF3-5F	GAGCTTTTCCCCTAGGCATC
IKZF3-5R	ACACTGGGCTCTGAGGAATG
IKZF3-6F	AGTTGTCACAGAGCCCCAAT
IKZF3-6R	ATCTGTCTGCCCCCAGTAAA
IKZF3-7F	GCATCCAACTTCGGACACTT
IKZF3-7R	TTTAACAGAGGTTAAGCTAGGAAAGG
IKZF3-8F	CTTGGCAGTGTTCCCTTTGT
IKZF3-8R	GGCGAGGTCATTGGTTTTTA
TP53-2-3F	CCATTCTTTTCCTGCTCCAC
TP53-2-3R	TCAAATCATCCATTGCTTGG
TP53-4F	CCTGGTCCTCTGACTGCTCT
TP53-4R	GACAGGAAGCCAAAGGGTGA
TP53-5F	GTTTCTTTGCTGCCGTCTTC
TP53-5R	GAGCAATCAGTGAGGAATCAGA
TP53-6F	CTGCTCAGATAGCGATGGTG
TP53-6R	TTGCACATCTCATGGGGTTA
TP53-7F	GCACTGGCCTCATCTTGG
TP53-7R	GGGATGTGATGAGAGGTGGA
TP53-8F	TCTGGCTTTGGGACCTCTTA
TP53-8R	GGAAAGAGGCAAGGAAAGGT
TP53-9F	AAGCAAGCAGGACAAGAAGC
TP53-9R	TGTCTTTGAGGCATCACTGC
TP53-10F	AACTTGAACCATCTTTTAACTCAGG
TP53-10R	GAAGGCAGGATGAGAATGGA
TP53-11F	AAAGCATTGGTCAGGGAAAA
TP53-11R	GCAAGCAAGGGTTCAAAGAC

3.11 Statistical analysis

Overall survival (OS) was calculated taking into account the first date of diagnosis of BPDCN to death of disease or to last follow-up. Kaplan-Meier method was used to estimate the distribution of OS and differences in survival between groups were assessed using the log rank test with the SPSS v.19 software (SPSS Inc, Chicago, IL, USA). A P value of $\leq .05$ was considered statistically significant.

CHAPTER 4: Results and Discussion

4.1 Part I: Whole-exome sequencing of chronic myeloid leukemia progression

The mechanisms of transformation from chronic-phase (CP) to blast-crisis (BC) of chronic myeloid leukemia (CML) are heterogeneous and poorly understood [96]. The most frequently observed genetic aberrations at this advanced stage include a second Philadelphia chromosome (Ph), trisomy 8, isochromosome 17, and trisomy 19, alone or in various combinations [97]. Clonal cytogenetic evolution appears to be the most consitent predictor of blastic transformation, present in up to 80% of patients [98].

At present, little is known about the mutational profiling of the different steps of CML progression. High- throughput genomic studies have been partially reported and only few mutations present at BC have been described [99]. The most common mutations in myeloid BC involve TP53 *and* RUNX1*, found in about 25% and 40% of the cases, respectively [100, 101]. In lymphoid BC, the most commonly reported mutations involve* CDKN2A *[102],* IKZF1 *[103] and* RB1 *[104] in 50%, 55% and 20% of cases, respectively. In addition, clinical data indicate that development of ABL1 mutations during treatment and/or disease progression are associated with a poorer outcome [105-108].*

Although these findings have identified mutations at single time-points, the genomic basis of the progression of CML over time from diagnosis to end-stage disease remains unclear. Here, we used whole-exome sequencing (WES) to analyse the exome of one individual at three different time-points of CML progression. We also included additional patient samples to explore and define the incidence and recurrence of some mutations found in CP and BC. This approach allowed us to identify, even in the CP phase, mutations with prognostic and predictive significance as well as alterations that have not been described so far in this disease.

4.1.1 Case report

The patient was a 65-year-old men diagnosed with a Ph+ CML and presented with a hypercellular bone marrow containing 2% blast cells. Karyotype analysis showed the Ph chromosome in all analyzed metaphases at diagnosis. The patient was treated with Imatinib® (400 mg/day) achieving, along the treatment, complete hematological response after one month and complete cytogenetic response after 12 months. Conversely, he never achieved major molecular response during this period. Unfortunately, after 14 months from diagnosis the patient progressed to a myeloid blast crisis that did not respond to a 2nd line treatment (Dasatinib® + Idarubicin-AraC) and the patient died of the disease 18 months after diagnosis. At that time, the G-banding analysis showed a complex karyotype (Figure 5).

4.1.2 Identification of candidate somatic mutations in CML progression

To identify somatic mutations during CML progression, we performed an exome sequencing of this index patient at three different phases of the disease: chronic phase (CP), complete cytogenetic response (CCyR) and in disease progression (blast crisis, BC). The preparation of shotgun libraries from the leukemic (CP, CCyR and BC) and non-leukemic genomic DNA obtained from the patient, followed by in-solution exome capture, was performed with the use of a commercial platform (Agilent) covering 50 Mb of coding exons (approximately 1.60% of the genome). After massively parallel sequencing with the Genome Analyzer IIx (Illumina), candidate somatic mutations were identified using the pipeline detailed at Material and Methods chapter of this thesis (pg.34).

After discarding the variants present in the matched normal DNA and in the dbSNP132 database, we obtained a total of 3123, 7678 and 3306 single nucleotide substitutions (SNSs) and small insertions and deletions (indels) for CP, CCyR and BC, respectively. Next, we selected only those variants within coding regions that, passing depth and quality controls, were frameshift, stop gain/loss and non-synonymous amino acid changes predicted to produce deleterious damage in the protein structure. Finally, this resulted in 13, 7 and 15 SNSs/indels for CP, CCyR and BC, respectively, affecting a total of 20 genes (Table 7 and Figure 6A).

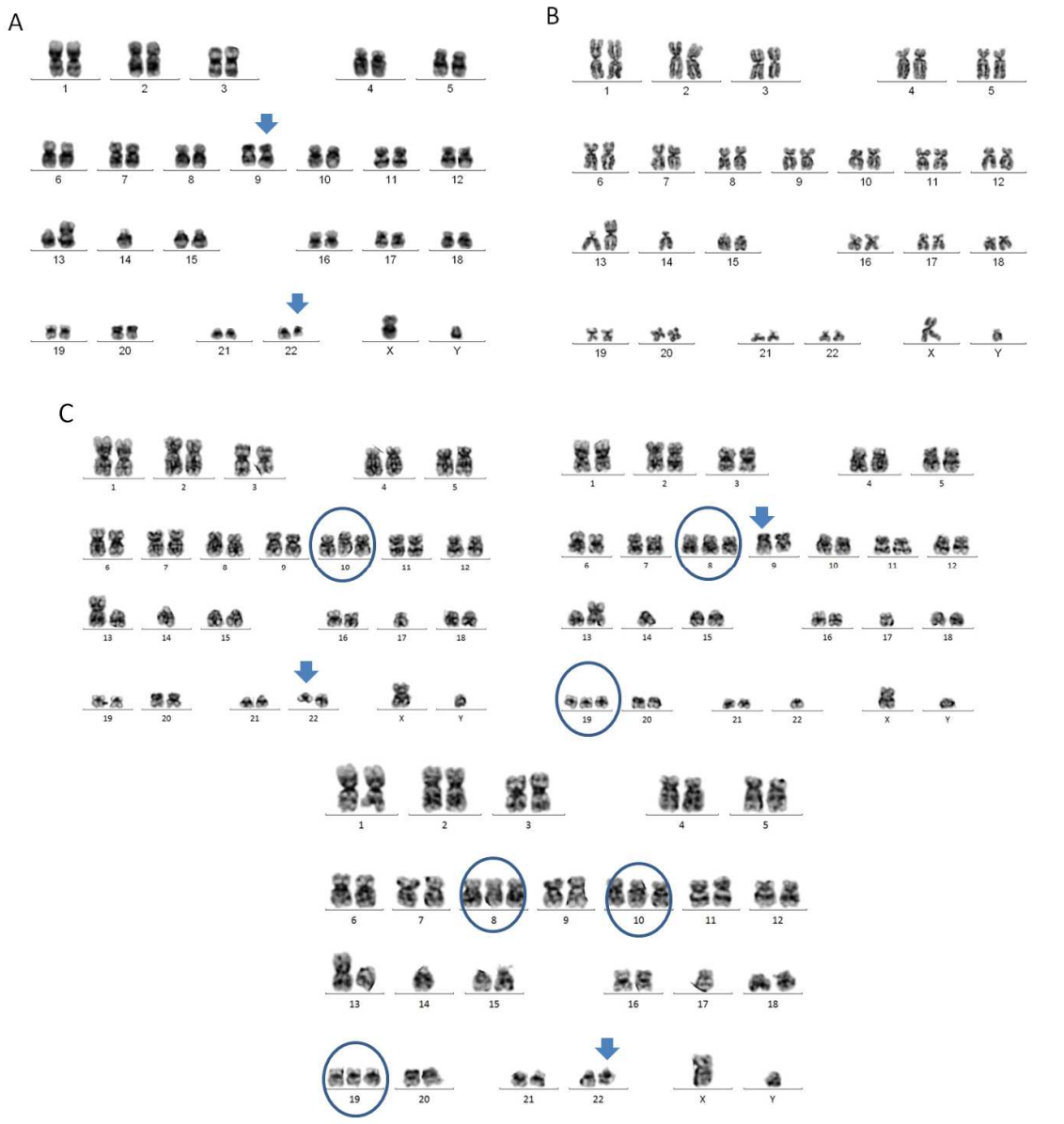

Figure 5: Karyotype during CML progression (A) chronic phase: 45, XY, t(9;22)(q34;q11.2), rob(13;14)(q10;q10) [20]; (B) complete cytogenetic response: 45, XY, rob(13;14)(q10;q10) [20]; and (C) blast crisis: 44, XY, t(9;22)(q34;q11.2), rob(13;14)(q10;q10), -17 [5]; 45, XY, t(9;22)(q34;q11.2), +10, rob(13;14)(q10;q10), -17 [5]; 46, XY, t(9;22)(q34;q11.2), +8, +10, rob(13;14)(q10;q10), -17 [3]; 47, XY, t(9;22)(q34;q11.2), +8, +10, rob(13;14)(q10;q10), -17, +19 [2]; 45, XY, rob(13;14)(q10;q10) [5]. The blue arrows indicate the t(9;22) and the blue circles, the additional chromosome alterations.

Table 7: SNSs/indels filtering steps of CML WES

Filter	CP	CCyR	BC
Total	3123	7678	3306
Somatic	719	1839	869
Exonic	372	878	412
Consequence*	77	81	66
Quality Control	13	7	15

*Frameshift, stop gain and non-synonymous-deleterious

Among those SNSs and indels, we selected the variants present with a frequency higher than 20%, due to the limitations of Sanger sequencing (Table 8). We validated mutations in genes known to be involved in CML (such as *ASXL1* and *TP53*) [109, 110] as well as in genes that have not been described so far in this disease, such as *UBE2G2*, *ZEB2* and *IKZF3* (Figure 6B).

Interestingly, *TP53* mutation (p.G244S) is present in all the stages which lead to suggest its potential role as a tumor-initiating events. On the other hand, *ASXL1* (p.G679*), *UBE2G2* (p.D35V), *ZEB2* (p.L420R) and *IKZF3* (p.E272K) were present in both CP and BC, also suggesting that these alterations are initiating events and co-occur with BCR-ABL1 fusion gene (Figure 6A). All theses SNSs were validated using polymerase-chain-reaction (PCR) amplification and direct DNA sequencing of the same samples that were subjected to WES. Finally, we also found SNSs unique to the BC sample (*ARSD* and *AADACL3* mutations) that may, therefore, contribute to leukemic transformation. However, we were not able to validate these two mutations probably due to Sanger sequencing limitations.

Among the genes previously reported in CML progression [99, 111-114], we detected *ASXL1* and *TP53* mutations. Interestingly, *TP53* mutation (p.G244S) was found in the three phases of CML progression and is present in the Catalogue of Somatic Mutations In Cancer (COSMIC) database [115]. There are some controversial results regarding its role in blastic transformation. Whereas some authors have observed mutations in 25%-30% of BC [97], a recent work by Grossman *et al*. [99], in which deep sequencing approach was used, they found a *TP53* mutation only in one out of 39 (2,6%) BC patients analyzed.

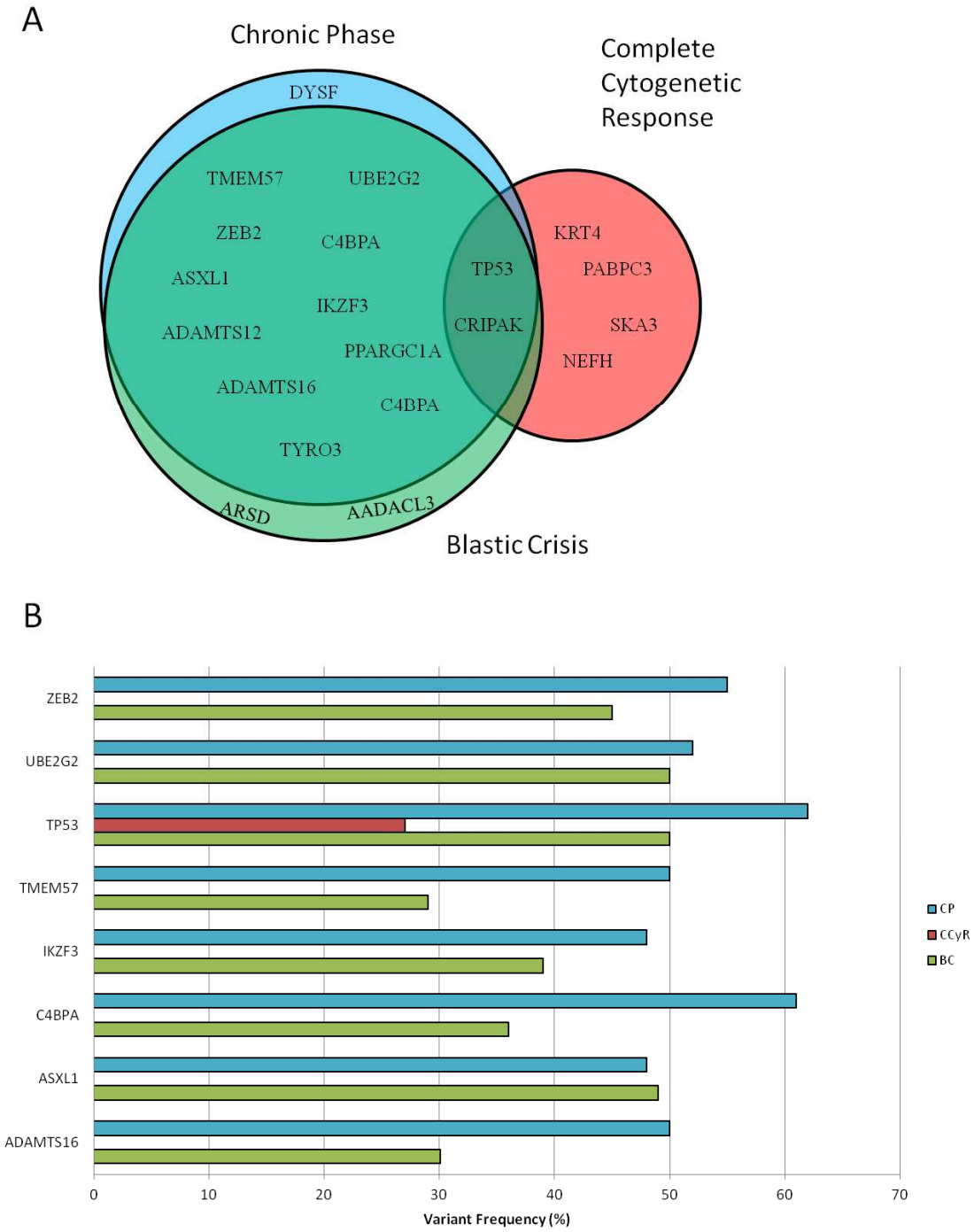

Figure 6: Mutational profile by whole exome sequencing along the three phases of CML. (A) Venn diagram representing the distribution of mutated genes distribution in CML progression; (B) Frequencies of the mutant alleles of every affected gene at the three time-points of CML progression. Sanger validation was conducted for those genes with frequencies above 20%: ZEB2, UBE2G2, TP53, TMEM57, IKZF3, C4BPA, ASXL1 and ADAMTS16.

Table 8: Somatic SNSs and indels in CML progression

CML-Phase	Position	Gene Name	QUAL	Coverage	Freq_Var(%)	Consequence	DNA Level (cDNA)	Protein Level	PolyPhen	SIFT
CP	5_33549323_G/T	ADAMTS12	503,94	24	8	NON_SYNONYMOUS_CODING	NM_030955.2:c.4291C>A	p.Pro1431Thr	probably_damaging	tolerated
CP	5_5232589_G/A	ADAMTS16	1206,77	96	50	NON_SYNONYMOUS_CODING	NM_139056.2:c.1810G>A	p.Gly604Arg	probably_damaging	deleterious
CP	20_31022550_G/T	ASXL1	323,4	27	48	STOP_GAINED	NM_015338.5:c.2035G>T	p.Gly679*	NA	NA
CP	1_207305072_C/G	C4BPA	1060,02	57	61	STOP_GAINED	NM_000715.3:c.1071C>G	p.Tyr357*	NA	NA
CP	4_1388623_-/CA	CRIPAK	488,27	122	9	FRAMESHIFT_CODING	NM_175918.3:c.324_325insCA	p.Thr109Glnfs*82	NA	NA
CP	2_71801335_-/AGGCGG	DYSF	320,44	24	20	FRAMESHIFT_CODING	NM_001130987.1:c.3236delinsAGGCGG	p.Glu1081Glyfs*59	NA	NA
CP	17_37922621_C/T	IKZF3	707,81	50	48	NON_SYNONYMOUS_CODING	NM_012481.3:c.952G>A	p.Glu318Lys	probably_damaging	deleterious
CP	4_23825925_23825926delinsCT	PPARGC1A	999,13	113	35	NON_SYNONYMOUS_CODING	NM_013261.3:c.854_855delinsAG	p.Ser285Lys	probably_damaging	deleterious
CP	1_25824890_A/G	TMEM57	2076,54	155	50	NON_SYNONYMOUS_CODING	NM_018202.4:c.1928A>G	p.Lys643Arg	probably_damaging	tolerated
CP	17_7577551_C/T	TP53	768,45	47	62	NON_SYNONYMOUS_CODING	NM_000546.4:c.730G>A	p.Gly244Ser	probably_damaging	deleterious
CP	15_41870343_C/T	TYRO3	244,25	23	4	NON_SYNONYMOUS_CODING	NM_006293.3:c.2542C>T	p.Arg848Trp	probably_damaging	deleterious
CP	21_46197270_T/A	UBE2G2	1222,13	87	52	NON_SYNONYMOUS_CODING	NM_003343.5:c.188A>T	p.Asp63Val	probably_damaging	deleterious
CP	2_145157495_A/C	ZEB2	954,8	67	55	NON_SYNONYMOUS_CODING	NM_014795.3:c.1259T>G	p.Leu420Arg	probably_damaging	deleterious
CCyR	4_1388623_-/CA	CRIPAK	427,91	141	9	FRAMESHIFT_CODING	NM_175918.3:c.324_325insCA	p.Thr109Glnfs*82	NA	NA
CCyR	12_53207608_-/CAG	KRT4	1437,89	198	5	IN-FRAME	NM_002272.2:c.457_458insCAG	p.Gly152_Gly153insAla	NA	NA
CCyR	22_29885623_-/AGGAAG	NEFH	894,64	248	4	IN-FRAME	NM_021076.3:c.1994_1995insAGGAAG	p.Lys665_Ala666insGlyArg	NA	NA
CCyR	5_43613188_TGCCCAAATCCAA/-	NNT	264,02	91	18	FRAMESHIFT_CODING	NM_012343.3:c.330_342del	p.Ala111Glyfs*23	NA	NA
CCyR	13_25671311_TATGA/-	PABPC3	225,27	197	9	FRAMESHIFT_CODING	NM_030979.2:c.976_980del	p.Met326Glyfs*21	NA	NA
CCyR	13_21729832_-/TGGAATTT	SKA3	1464,5	81	9	SPLICE_SITE,FRAMESHIFT_CODING	NM_145061.5:c.1238_1238+1insTGGAATTT	p.*413Cysfs*4	NA	NA
CCyR	17_7577551_C/T	TP53	250,69	45	27	NON_SYNONYMOUS_CODING	NM_000546.4:c.730G>A	p.Gly244Ser	probably_damaging	deleterious
BC	1_12785494_G/-	AADACL3	575,54	187	5	FRAMESHIFT_CODING	NM_001103170.1:c.584del	p.Cys195Phefs*17	NA	NA
BC	5_33549323_G/T	ADAMTS12	633,95	46	50	NON_SYNONYMOUS_CODING	NM_030955.2:c.4291C>A	p.Pro1431Thr	possibly_damagin	tolerated
BC	5_5232589_G/A	ADAMTS16	723,43	103	30	NON_SYNONYMOUS_CODING	NM_139056.2:c.1810G>A	p.Gly604Arg	possibly_damagin	deleterious
BC	X_2835999_CCACGCCGG/-	ARSD	278,35	29	17	IN-FRAME	NM_001669.2:c.701_709del	p.Ala234_Val236del	NA	NA
BC	20_31022550_G/T	ASXL1	390,7	35	49	STOP_GAINED	NM_015338.5:c.2035G>T	p.Gly679*	NA	NA
BC	1_207305072_C/G	C4BPA	511,16	53	36	STOP_GAINED	NM_000715.3:c.1071C>G	p.Tyr357*	NA	NA
BC	4_1388623_-/CA	CRIPAK	501,39	95	14	FRAMESHIFT_CODING	NM_175918.3:c.324_325insCA	p.Thr109Glnfs*82	NA	NA
BC	3_75714825_G/-	FRG2C	294,25	243	20	FRAMESHIFT_CODING	NM_001124759.1:c.483del	p.Arg161Serfs*5	NA	NA
BC	17_37922621_C/T	IKZF3	618,13	57	39	NON_SYNONYMOUS_CODING	NM_012481.3:c.952G>A	p.Glu318Lys	possibly_damagin	deleterious
BC	4_23825925_23825926delinsCT	PPARGC1A	1318,91	117	41	NON_SYNONYMOUS_CODING	NM_013261.3:c.854_855delinsAG	p.Ser285Lys	possibly_damagin	deleterious
BC	1_25824890_A/G	TMEM57	1452,52	199	29	NON_SYNONYMOUS_CODING	NM_018202.4:c.1928A>G	p.Lys643Arg	possibly_damagin	tolerated
BC	17_7577551_C/T	TP53	228,95	17	50	NON_SYNONYMOUS_CODING	NM_000546.4:c.730G>A	p.Gly244Ser	possibly_damagin	deleterious
BC	15_41870343_C/T	TYRO3	142,13	38	20	NON_SYNONYMOUS_CODING	NM_006293.3:c.2542C>T	p.Arg848Trp	possibly_damagin	deleterious
BC	21_46197270_T/A	UBE2G2	1439,23	104	50	NON_SYNONYMOUS_CODING	NM_003343.5:c.188A>T	p.Asp63Val	possibly_damagin	deleterious
BC	2_145157495_A/C	ZEB2	808,39	69	45	NON_SYNONYMOUS_CODING	NM_014795.3:c.1259T>G	p.Leu420Arg	possibly_damagin	deleterious

4.1.3 Defining *TP53*, *IKZF3* and *ASXL1* mutation frequencies in CML diagnosis and progression

Due to their biological and clinical relevance, we analyzed the *TP53, ASXL1* and *IKZF3* gene in an independent set of 26 samples consisting of 13 paired CML and BC/NCgR (no cytogenetics response) samples. No mutations in *TP53* were found. Although it seems that mutations affecting *TP53* gene can be considered as a rare event in the progression of CML, the presence of the mutation in all three phases confirm that it is clearly related to a non standard response to the treatment and/or clinical progression.

Regarding *ASXL1*, mutations in this gene during CML evolution were described for the first time by Boultwood *et al*. [112] They found mutation in six out of 41 patients analyzed (3 CP and 3 BC non paired samples; 14,5%), however in the only BC sample for which CP paired sample had available, they did not detect the same mutation. Subsequently, other authors have observed *ASXL1* mutations in similar frequencies [99, 116, 117] and among these earlier works, *ASXL1* gene is the commonly affected gene. According with these previous findings, we detect a missense mutation in the patient sequenced and we also observed two different mutations in four out of 26 (15%) in BC/NCgR and also in the CP paired samples analyzed (Table 9). All these data together corroborate that *ASXL1* mutations constitute an early event and might cooperate with other alteration such *BCR-ABL1*.

Regarding mutations found in CP and BC, it is important to mention the mutation observed in *IKZF3* (p.E272K). *IKZF3* is a member of the Ikaros transcription factors family, which are important regulators of lymphoid differentiation. Among the five members of the family, *IKZF1* have been found deleted and mutated in acute lymphoblastic leukemia (ALL), CML blastic phase and *BCR-ABL1* positive ALL suggesting a pathogenetic contribution to leukemic transformation [118]. In addition, this genetic alteration have been found also in 19% of patients with blast-phase myeloproliferative neoplasm leading to suggest a potential pathogenic role in the myeloid lineage [119]. To date, mutations in other members apart from *IKZF1* have never been described in human leukemia. In order to explore this genetic feature, we screened for mutation in *IKZF3* gene in the 26 samples using conventional PCR and Sanger sequencing. *IKZF3*

were also mutated in 8% of CP and BC/NCgR samples, suggesting a potential role of this gene in myeloid leukemia (Table 9).

4.1.4 Clonality assessment in CML progression

The application of WES approach allowed determining the clonality and clonal evolution patterns during the CML progression. In this sense, 93% of the selected SNSs that were present in the CP were also seen in BC. In fact, the percentages of reads of the mutant alleles identified for the most relevant genes were the same (around 50%) both at CP and at BC (Fig.6B). As expected, these data suggest that the same clone was present in CP and BC.

In summary, WES allowed the identification of a large number of mutated genes, even at the chronic phase of CML that harbor prognostic and predictive significance, such as *ASXL1* and *TP53*, both found mutated at CP and BC. The study of the mutation profile through the course of the disease indicated that, at least in this patient, the number and the type of mutations were similar at CP and BC. In addition, we identified for the first time deleterious mutations in *IKZF3*, *UBE2G2* and *ZEB2* in CML at diagnosis. While current diagnostic procedures recommend the study of *ABL1* mutations in non-responders patients, our data suggest that sequencing a wider panel of genes could be also beneficial in the clinical management of these patients.

Table 9: Chromosomal and molecular features in 13 CP and BC/NCgR paired samples

Id	Type	Age (years)	Follow-up (months)	Karyotype	Mutational status		
					ASXL1	*IKZF3*	*TP53*
1*	CP	65/M	18	45,XY,t(9;22)(q34;q11.2),rob(13;14)(q10;q10)c[20]	c.2035G>T	c.952G>A	c.730G>A
	BC			44,XY,t(9;22)(q34;q11.2),rob(13;14)(q10;q10)c,-17 [19]/46,XY,t(9;22)(q34;q11.2),+10,rob(13;14)(q10;q10)c,-17[1]	c.2035G>T	c.952G>A	c.730G>A
2	CP	31/F	6	46,XX,t(9;22)(q34;q11.2)[20]	c.2498_2501del	wt	wt
	NCgR			46,XX, t(9;22)(q34;q11.2)[20]	c.2498_2501del	wt	wt
3	CP	70/M	3	46,XY,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	NCgR			46,XY,t(9;22)(q34;q11.2)[6]/46,XY[14]	wt	wt	wt
4	CP	41/F	5	46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	BC			47,X,-X,+der(1)t(1;?)(p12;?),-5,+6,+8,t(9;22)(q34;q11.2),-9,-14,-17,-18,+21,+der(22)t(9:22)(q34;q11.2),+1mar[cp2]/46,XX[6]	wt	wt	wt
5	CP	54/M	5	46,XY,t(2;13)(p36;?),t(9;22)(q34;q11)[20]	wt	wt	wt
	NCgR			46,XY,t(2;13)(p36;?),t(9;22)(q34;q11)[1]/46,XY[19]	wt	wt	wt
6	CP	85/M	7	46,XY,t(9;22)(q34;q11.2)[4]/46,XY,t(9;9;22)(p12;q34;q11.2)[10]/46,XY[6]	wt	wt	wt
	NCgR			46,XY[11]/46,XY,t(9;9;22)(p12;q34;q11.2)[5]	wt	wt	wt
7	CP	36//M	58	46,XY,t(9;22)(q34;q11.2)[7]/46,XY[8]	wt	wt	wt
	NCgR			46,XY,t(9;22)(q34;q11.2)[2]/46,XY[18]	wt	wt	wt
8	CP	34/M	12	46,XY,t(9;22)(q34;q11.2),der(3)t(3;?)(?;?)[7]/46,XY,t(9;22)(q34;q11.2)[8]	wt	c.71C>T	wt
	BC			46,XY,t(9;22)(q34;q11.2)[2]/46,XY[26]	wt	c.71C>T	wt
9	CP	8/M	4	46,XY[3]/46,XY,t(9;22)(q34;q11.2)[17]	c.2598A>G	wt	wt
	BC			46,XY[2]/47,XY,+X,-7,der(9)i(9)(q34;q11.2),+mar[18]	c.2598A>G	wt	wt
10	CP	44/F	24	46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	NCgR			46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
11	CP	73/F	20	46,XX,t(9;22)(q34;q11.2)[20]	c.1937dup	wt	wt
	NCgR			46,XX,t(9;22)(q34;q11.2)[20]	c.1937dup	wt	wt
12	CP	60/F	14	46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	BC			47,XX,t(9;22)(q34;q11.2),+t(9;22)(q34;q11.2)[10]/46,XX,t(9;22)(q34;q11.2)[8]/46,XX[2]	wt	wt	wt
13	CP	69/F	3	46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	NCgR			46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
14	CP	NA	24	46,XX,t(9;22)(q34;q11.2)[20]	wt	wt	wt
	BC			45,XX,-7,t(9;22)(q34;q11.2)[8]/46,XX[12]	wt	wt	wt

4.2 Part II: Genomic characterization of chronic neutrophilic leukemia

Chronic neutrophilic leukaemia (CNL) is a rare but distinct myeloproliferative neoplasm (MPN) characterized by the presence of leukocytosis, hypercellularity of bone marrow (BM), a sustained mature neutrophilia in peripheral blood (PB), hepatosplenomegaly, lack of detectable BCR-ABL1 transcripts, as well as by the absence of rearrangements in PDGFRA/B *and* FGFR1. *According 2008 WHO classification, the diagnosis is based on the neoplastic expansion of granulocytic cells and the exclusions of genetic drivers that are known to occur in other MPN [21].*

Recent studies identified the presence of CSF3R *mutations in more than 50% of patients with CNL or atypical chronic myeloid leukemia (aCML).* CSF3R *encodes the trans-membrane receptor for granulocyte colony-stimulating factor (G-CSF; CSF3), which provides the proliferative and survival signal for granulocytes and also contributes to their differentiation and function [120]. The oncogenic* CSF3R *mutations are molecular markers of sensitivity to inhibitors of SRC family-TNK2 and JAK kinases and may provide a new avenue for therapy [38, 121].*

However, no comprehensive massive next-generation sequencing study has been performed to explore the whole picture of the disease. Here we use whole-exome sequencing (WES) and RNA sequencing (RNA-seq) to look at additional genetic abnormality in CNL. In addition to the CSF3R *p.Thr618Ile mutation, we were able to identify others mutations (on spliceosome, chromatin remodeling and methylation genes), loss of heterozygozity (LOH), splicing abnormalities and fusion genes that, together, could explain the complexity and aggressiveness of this disease.*

4.2.1 Case Report

A 66-year-old man was diagnosed with CNL, according to 2008 WHO classification. At presentation, the patient presented peripheral blood leukocytosis ($66x10^9$/L), segmentted neutrophils and band forms were 91.5% of the white blood cells counts (WBCs), immature granulocytes were <10% of WBCs and myeloblasts were <1%. The aspirate showed a hypercellular BM with neutrophilic granulocytes increased in number and a percentage and myeloblasts 0.5% of WBCs. No dysplastic features were observed in the myeloid lineages. His Zubrod Performance Status (ECOG) was 1. G-banding chromosome analysis revealed a normal karyotype (46,XY[20]) and molecular biology studies were negative for BCR-ABL1 transcripts and JAK2 V617F mutation. The patient was treated with hydroxyurea, but, unfortunately, died seven months after diagnosis due to an intensification of the disease.

4.2.2 Mutations in coding sequencing by WES of CNL

To improve our understanding of the genes involved in the pathogenesis of CNL, WES was performed on matched tumor and normal sample from the patient. Candidate somatic mutations were identified using RUbioSeq software [89]. The bioinformatics analysis and the filtering steps to identify the coding variants are detailed in the Material and Methods chapter of this thesis (pg.34). In total, we found 1437 somatic mutations and, among them, 797 were intronic, intergenic, affecting non-conding-RNA or untranscribed regions. From the 640 exonic variants, we selected only those variants within coding regions that, after passing sequencing depth and quality filters were, frameshift, stop gain/loss and non-synonymous amino acid changes predicted to produce a deleterious effect in the protein structure.This process resulted identified a total of 24 SNSs/indels. In addition to the *CSF3R* p.Thr618Ile mutation, we found by WES sequencing mutations in: *U2AF1*, *TET2*, *LUC7L2* and *ASXL1* (Figure 7 and Table 10).

The current study has several focus of interest, firstly, it confirms the observations by Maxon *et al.* [38] and Pardanani et al [121], regarding the association between CNL and *CSF3R* mutations. Secondly, it presents a complete picture of the mutational profiling of CNL, certainly more complex than expected from these

previous reports. In fact, we found and validated mutations affecting both splicing machinery and epigenetic genes as mentioned. Kosmider *et. al.* [122], very recently, showed that *CSF3R* somatic mutations can be identified in ~4% of chronic myelomonocytic leukemias. These mutations, which affect distinct residues in *CSF3R* as compared to CNL, are frequently associated with mutations in *ASXL1* gene and have a poor prognostic impact on overall and acute myeloid leukemia (AML)-free survival. Taking into consideration these reported data, it seems that also the CNL genome displays a combination of few mutations with a pattern of cooperation with strong biologic relationship among genes and categories, being this scenario similar to the one observed in others myeloid disorders, such as AML [69]. In a very similar way to what happens in other myeloid disorders, this type of cooperating mutations in epigenetic genes was also found in our CNL patient, whose leukemic genome showed mutated copies of *ASXL1* and *TET2* genes.

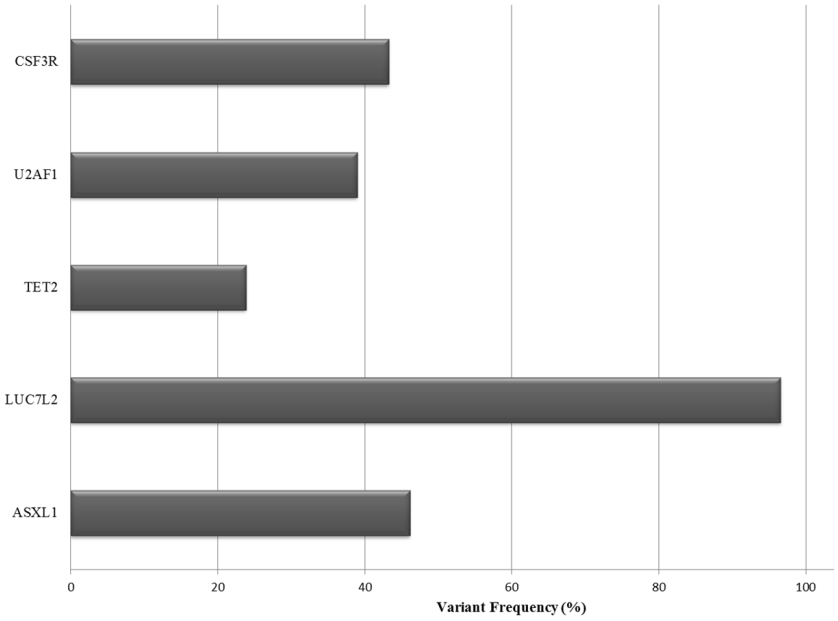

Figure 7: Mutational profiling of CNL. In addition to the *CSF3R* p.Thr618Ile mutation, WES revealed mutations in: *U2AF1*, *TET2*, *LUC7L2* and *ASXL1*. *LUC7L2* mutation was found in homozygosis.

Table 10: Description of validated mutations in the CNL patient

Chrom	Position	Gene Name	Zygosity	Ref	Variant	Var Freq	Coverage	AA change	Type	Provean	SIFT
20	31024563	ASXL1	Het	C	T	46,12	2433	Q1350*	Stop Gain	NA	NA
7	139060825	LUC7L2	Hom	C	T	96,55	261	R93*	Stop Gain	NA	NA
4	106157539	TET2	Het	CG	C	23,81	126	NA	Frameshift	NA	NA
21	44514777	U2AF1	Het	T	C	38,96	2477	Q157R	Single AA Change	Deleterious	Damaging
1	36933434	CSF3R	Het	G	A	43,23	1203	T618I	Single AA Change	Deleterious	Damaging

NA: no aplicable; Het: heterozygotic; Hom: homozygotic; AA: aminoacid

4.2.3 Loss of heterozygosity in 7q and detection of *LUC7L2* expression

The variant allelic frequency of *LUC7L2* mutation was found to be high (more than 95%). Makishima *et. al.* [123] previously described mutations of this gene in MDS and, since this gene is located in the 7q region, a frequently lost chromosomal region in myeloid leukemias [124], we decided to investigate whether a critical deletion or a loss of heterozygosity (LOH) affecting this genomic region was also present in the patient. To study this phenomenon, we interrogate our WES data for LOH across the genome of the sample. Interestingly, we found a LOH of 53.2MB in chromosome 7q including the locus of the *LUC7L2* gene. Allele frequencies of each SNP along chromosome 7 are shown in Figure 8. Since no del(7q) was detected with metaphase cytogenetics, our study demonstrates for the first time mutations in *LUC7L2* accompanied by a copy-neutral LOH (uniparental disomy) in 7q in a patient with an aggressive CNL phenotype.

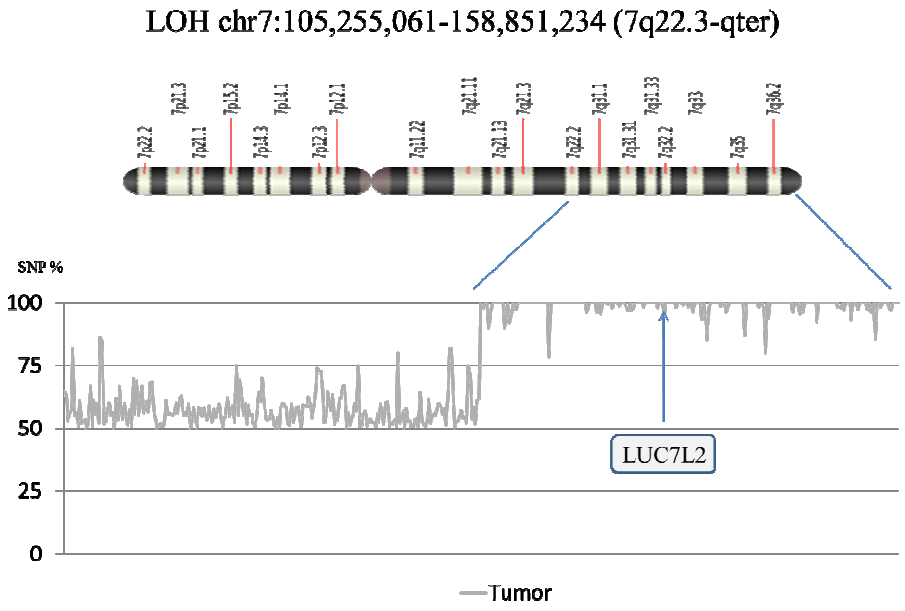

Figure 8: 7q LOH in CNL. LOH of 53.2MB in chromosome 7q including the *locus* of *LUC7L2*

To evaluate the biological consequences of this homozygous mutation, we explored *LUC7L2* expression in the bone marrow cells of the patient and in some myeloid leukemia cell lines (Figure 9). By real-time PCR, we observed a total abrogation of the expression of *LUC7L2* in the patient cells as well as general down regulation in myeloid leukemia cell lines. This downregulation of *LUC7L2* expression is a completely novel finding in the field and its biological and clinical consequences still remain to be further explored.

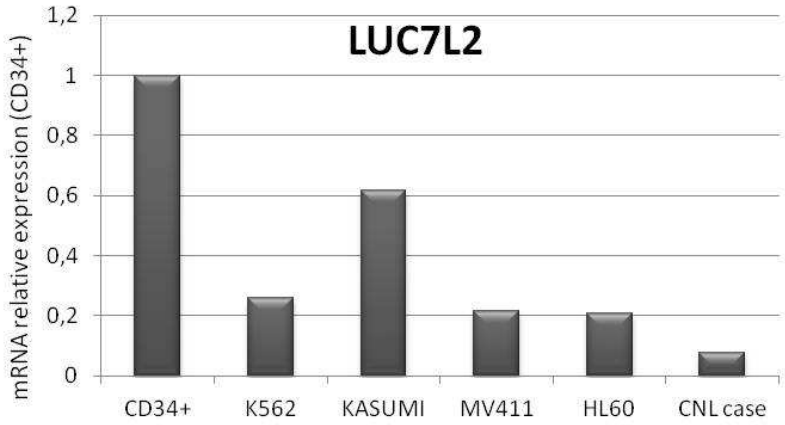

Figure 9: LUC7L2 expression in CNL patient. Abrogation of *LUC7L2* expression in CNL patient and down regulation in myeloid leukemia cell lines

4.2.4 Effects of mutations *in LUC7L2* and *U2AF1* on spliceosomal function

Conceptually, mutations of spliceosomal proteins, such as *LUC7L2* and *U2AF1,* could result in defective splicing, including intron retention, altered splice site recognition, or altered alternative splicing. To determine the functional consequences of these mutations in the proper mRNA production and splicing process in CNL, we performed a complete and massive sequencing of all mRNA (RNA-seq) present in the BM of our index patient as well as in CD34+ cells from a normal control. In the presence of functional spliceosomal machinery, sequencing reads are expected not to cross the intron/exon boundaries and therefore should not contain any intronic sequences. Although no clear genome-wide increase in intron retention was observed in the patient, as previously reported by Makishima *et. al.* [125] in some MDS cases, we found an altered pattern of splicing in the mRNA species transcribed from the *RUNX1* gene. At the 3' splice site of *RUNX1* intron 5, the un-spliced reads were almost

3 times more frequent in the mutated patient than the normal control (Figure 10). Because of the large numbers of diverse mutations in the splicing machinery, large studies will be needed to fully evaluate the impact of these mutations in splicing.

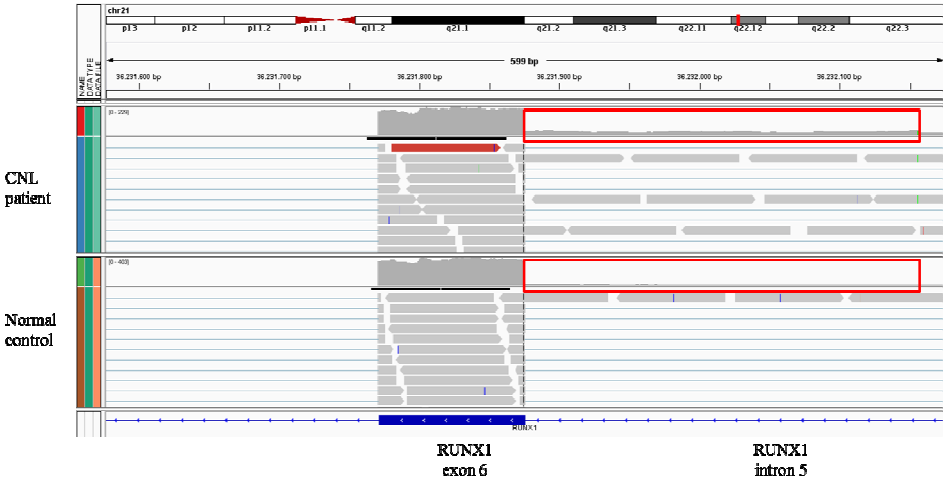

Figure 10: Integrative Genome Viewer image from RNA-seq data. We found an altered pattern of splicing in the mRNA species transcribed from the *RUNX1* gene. At the 3' splice site of *RUNX1* intron 5, the un-spliced reads were almost 3 times more frequent in the mutated patient than the normal control.

4.2.5 Identification of PIM3/SCO2 fusion transcript

Another putative effect of mutations in spliceossome machinery is the appearance of chimeric RNAs formed, as in a reciprocal translocation, by the junction of two originally distinct genes. To assess this kind of instability, we used our RNA-seq data. In fact, we identified a chimeric transcript involving *PIM3* and *SCO2* genes (both were located on 22q13.33), that was the result of an intrachromosomal inversion of approximately 0.6Mb in chromosome 22 (Figure 11). The *PIM3* oncogene belongs to the Ser/Thr protein kinase family, and PIM subfamily. This gene is overexpressed in hematological and epithelial tumors and is associated with MYC co-expression. It plays a role in the regulation of signal transduction cascades, contributing to both cell proliferation and survival, and provides a selective advantage in tumorigenesis [126]. Interestingly, inhibition of PIM kinases by pim kinase inhibitors in Myc-induced lymphoma resulted in cell in cell death [127]. Functional studies are needed to elucidate the role of this fusion gene in leukemogenesis.

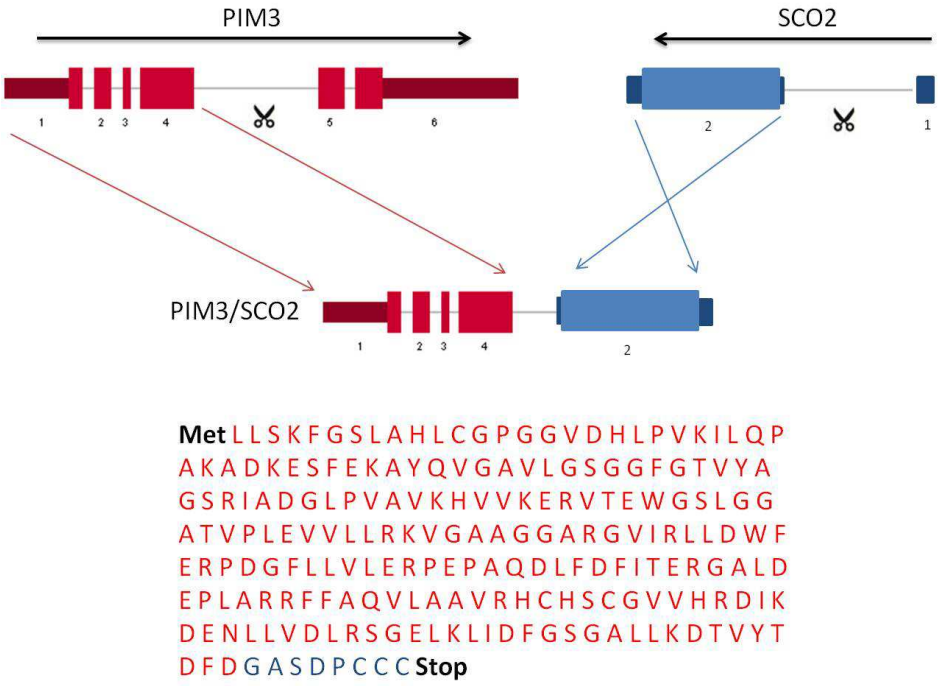

Figure 11: *PIM3/SCO2* fusion gene. *PIM3-SCO2* fusion gene result of an intrachromosomal inversion of approximately 0.6Mb in chromosome 22

4.2.6 Gene expression profiling

Finally, we also took advantage of the global expression profile that can be obtained from the RNA-seq data. We performed such analysis and identified the genes that were differentially expressed in the CNL cells compared with the normal CD34+ cells. We obtained 2022 genes upregulated and 1884 genes downregulated in CNL compared with the control (FDR=0.05). Interestingly, *IKZF1*, *CEBPA*, *PDGFRA/B*, *IDH1*, *MAPK1*, *TET2*, *CDK2*, *MEIS2/3*, and *PBX3* were upregulated in our CNL in comparison with control and, among the downregulated, we observed *IKZF2/3*, *HOX* genes, *UBE2G2*, *ASXL1*, *RUNX1/3*, *LUC7L2* and *YES1*.

To describe the effects of these deregulated expressions, we performed an Ingenuity Systems Analysis, detailed in the Material and methods of this thesis (pg. 35). Functional classification of genes differently expressed between CNL cells and normal control (CD34+) revealed an enrichment of categories like cell signaling; cell death and survival; and gene expression (B-H p-value <0.05) (Figure 12). These findings reconfirmed the significant pathophysiology of such several synergetic unique genetic defects in the general CNL cohort, as well as in our index case.

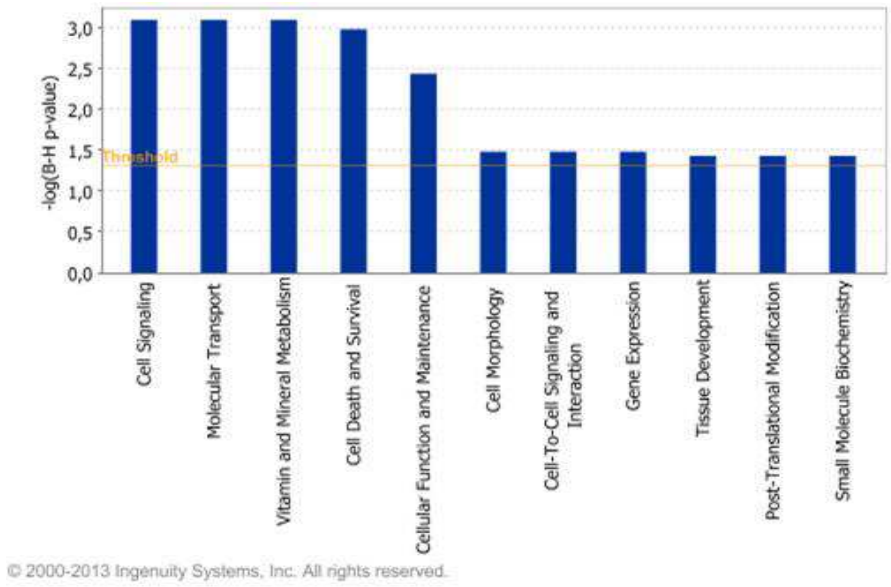

Figure 12: IPA classification in CNL. Functional classification of genes differently expressed between CNL cells and normal control (CD34+), according to IPA, revealed an enrichment of categories like cell signaling; cell death and survival; and gene expression (B-H p-value <0.05).

4.2.7 Clinical relevance of the mutation analysis

The oncogenic *CSF3R* mutations T618I strongly activate the JAK/STAT pathway and are sensitive to inhibitors of SRC family-TNK2 and JAK kinases and may provide a new avenue for therapy [128]. On the other hand, pim kinase inhibition could be a viable treatment strategy in certain human leukemias that rely on PIM3 kinase expression [127]. In addition, epigenetic modifiers provide new targets for therapeutic intervention and targeting these enzymatic activities are currently being explored from a therapeutic standpoint in several types of leukemia [129, 130]. Although, Pardanani *et al.* considering 35 cases of clinically suspected CNL, did not found alteration in the survival based on the presence or absence of *CSF3R* mutations, our reported patient had a rapid disease progression and died seven months after diagnosis, probably explained by the profoundly aberrant landscape of gene mutations and rearrangements with functional effects on the biology of the tumor cells.

In sum, our study provides, for the first time, a massive molecular and expression data, revealing a large amount of genomic alteration in CNL. In this complex scenario, a combination of new targets therapies may be considered as reasonable option for the therapeutic management of this aggressive and rare subtype of leukemia.

4.3 Part III: Blastic plasmacytoid dendritic cell neoplasm coding genome

Blastic plasmacytoid dendritic cell neoplasm (BPDCN), previously referred as "blastic NK-cell lymphoma/leukemia" or "agranular CD4+/CD56+ hematodermic neoplasm", is a rare and aggressive hematologic disease derived from precursors of a specialized subset of dendritic cells, and hence, after some discussion, is considered a myeloid-related disease according to the 2008 WHO classification of myeloid neoplasms [21, 41-46]. These tumor cells infiltrate skin, bone marrow (BM), peripheral blood (PB) and lymph nodes [47-50] and mainly affect elderly patients with the maximum incidence peaking at approximately 65 years of age.

Diagnosis is based on the immunophenotype of the blast cells that are characterized by the expression of CD4, CD43, CD56, CD123, BDCA-2/CD303, TCL1, and CTLA [51]. In addition, immunohistochemical markers, such as SPIB, BDCA-4, IRF-8, BCL11A and CD2AP, have been recently reported as tools for BPDCN diagnosis [51, 52]. There are no established genetic biomarkers that assist in the clinical management of BPDCN. Preliminary results from molecular studies suggest that abnormalities of genes known to confer a poor prognosis in other hematological malignancies, such as CDKN2A/CDKN2B, TET2 and TP53, are frequently seen in BPDCN [53, 54].

The overall prognosis for BPDCN is remarkably poor. Lymphoid-like chemotherapy is the preferred treatment, however, most patient relapse, resulting in a median overall survival (OS) of 12-14 months [55, 56]. Only high-dose therapy followed by allogeneic stem cell transplantation can provide durable control in this otherwise fatal condition [57, 58].

We designed a study to discover BPDCN-associated mutations based on sequencing the complete exome of 3 tumors, matched with normal samples. The discovered mutations detected were then further studied in an expanded cohort. These efforts led to the identification of several important disease associated mutations, the reconstruction of the genetic pathways underlying the BPDCN pathogenesis, and revealed associations between genetic events and clinically important features of the disease.

4.3.1 Mutations in coding sequencing by WES in BPDCN

To improve our understanding of the genes involved in the pathogenesis and clinical evolution of BPDCN, WES was performed on matched tumor and normal samples from 3 individuals with BPDCN. The bioinformatics analysis and the filtering steps to identify the coding variants are detailed in the Material and Methods chapter of this thesis (pg.34). In total, we found 509, 590 and 1505 somatic alterations in case 1, case 2 and case 3, respectively (Table 11). Among them, 347, 457 and 961 were intronic, respectively and therefore, excluded from further analysis. To gain insight into the genes affected by the leukemic process in BPDCN, we focused our research on the mutations that passed our quality control and were predicted to result in protein-coding changes, identifying a total of 347 mutations in 81 affected genes. The distribution of numbers and categories of somatically acquired exonic mutations and indels is represented in Figure 13. Unfortunately, when we crossed these 81 affected genes, we did not identify a common mutated gene between patients (Figure 14). Of interest, however, we identified, for the first time, deleterious mutations in *IKZF3* (*IKAROS* family zinc finger 3 - Aiolos), *HOXB9* (homeobox B9), *UBE2G2* (ubiquitin-conjugating enzyme E2G 2) and *ZEB2* (zinc finger E-box binding homeobox 2) in human leukemia.

Table 11: SNVs filtering steps of BPDCN WES

Filter	Case1	Case2	Case3
Total	3306	3839	9683
Somatic	509	590	1505
Exonic	162	133	544
Consequence*	63	59	240
Quality Control	28	8	49

*Frameshift, stop gain and non-synonymous-deleterious

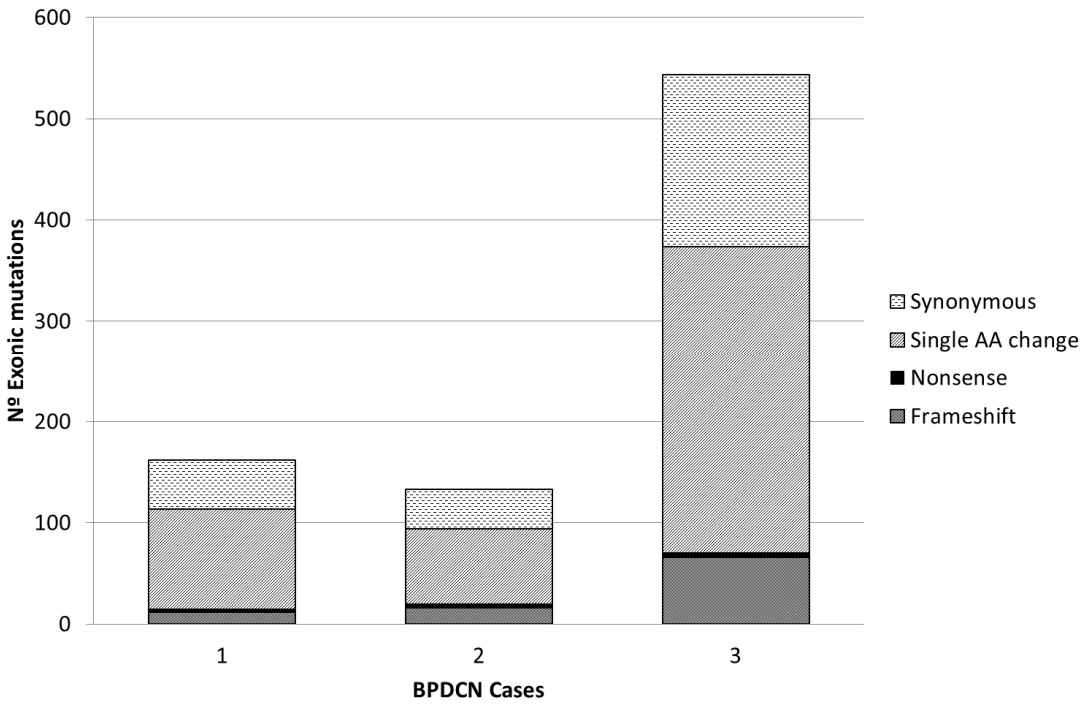

Figure 13: Distribution of the numbers and categories of somatically acquired exonic mutations and indels among the 3 cases studied by WES. In total, we found 509, 590 and 1505 somatic alterations in case 1, case 2 and case 3, respectively. Among them, 347, 457 and 961 are intronic, respectively.

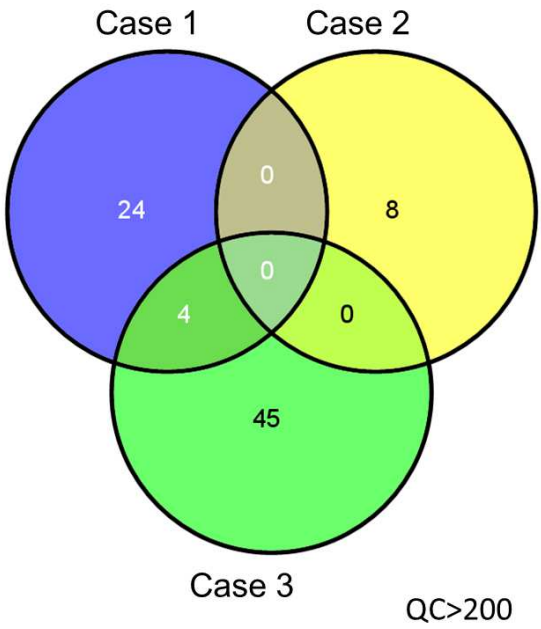

Figure 14: Venn diagram showing that there is no common mutated gene between patients (Exonic, frameshift, stop gain/loss and no synonymous mutations-deleterious; QC>200).

A functional clustering analysis of the mutated genes, using IPA software (Material and Methods pg. 35), showed a substantial enrichment of pathways involved in dermatological diseases and conditions, genetic disorder and hematological disease, among others (Figure 15). Looking in detail at the list of mutated genes within the hematological disease pathway, we observed that they have functions involved in DNA methylation (*TET1* and *TET2*), chromatin remodeling (*ASXL1*) and RNA splicing (*U2AF1*); all of which have previously been reported as mutated in myeloid neoplasms. These data suggested that BPDCN has a mutational profile similar to that of other well-defined myeloid leukemias.

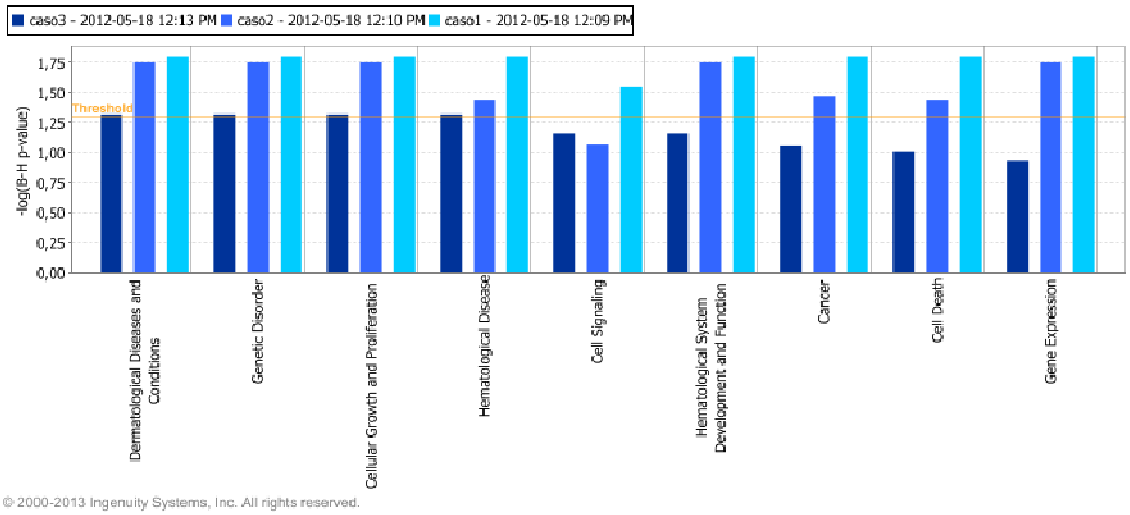

Figure 15: IPA classification in BPDCN. Substantial enrichment of pathways involved in dermatological diseases and conditions, genetic disorder, hematological disease, among others.

4.3.2 Screening for somatic mutations in regions recurrently altered in myeloid neoplasm by targeted next-generation sequencing

Taking the previous results into consideration, we decided to use a targeted re-sequencing approach to study these three pathways as well as transcription factors, protein kinase and ubiquitination genes involved in myeloid malignancies in a larger validation cohort of BPDCN samples. In detail, we applied a comprehensive and cost-effective target resequencing approach to indentify mutations in the coding sequences of 38 selected genes which are prone to harbor mutations involved in leukemogenesis (n=25) and have been identified in our previous analysis (n=13) (Table 5 on Material and Methods section). This panel is composed of genes from nine categories: DNA methylation (5 genes), chromatin remodeling (4 genes), transcription factors (10 genes), splicing machinery (5 genes), protein kinase activity (6 genes), ubiquitination (4 genes), RAS family (2 genes), nucleophosmin (1 gene) and tumor suppressor (1 gene). We included a total of 39 BPDCN tumour DNAs of which, after quality control for library preparation, 25 cases were suitable for analysis. On average, 98% of reads mapped uniquely to the target sequences, and we obtained 93% mean target coverage with an average read depth of 471 reads per base.

The bioinformatics analysis and the filtering steps to identify the coding variants are detailed in the Material and Methods chapter of this thesis (pg.36). After filtering out changes in the intronic regions and polymorphisms (present in dbSNP132), we identified a total of 141 variants in all sequenced exons of the 25 BPDCN patients, an average of 5.64 mutations per case. After discarding alterations in noncoding RNAs (ncRNAs), the 5' or 3' untranscribed regions (UTRs), and synonymous mutations, only 90 single nucleotid variants (SNVs) and indels resulted in an amino acid change, frameshift or stop codon (Table 12). We selected 52 SNVs/indels for validation by Sanger sequencing that showed frequencies above 20% and we were able to confirm 91% of the variants. In addition, with this resequencing strategy, we were able to detect all the alterations found in the three patients studied by WES, confirming the suitability of the panel for discovering genetic defects in the extension cohort.

Based on the pathogenic alterations, we observed mutations in 29 genes and a mean of 3.6 mutations per case. The extreme cases were one patient with 7 and

another two patients with no pathogenic alterations in the 38 studied genes (Figure 16A). *TET2* was the most frequently mutated gene (36%), followed by *ASXL1* (32%), *NRAS* (20%) and *NPM1* (20%), the *IKAROS* family (20%) and *ZEB2* (16%). In *TET2*, we identified ten heterozygous variants (4 frameshift, 4 nonsense and 2 missense changes) spread across the coding exons of the gene in nine patients. The frameshift p.Ile490Tyrfs*14 was found in two individuals and one patient had two different *TET2* mutations. *ASXL1* presented heterozygous mutations, mostly clustered in exon 12, in six patients. However, two mutations were found in exon 11. The majority of *ASXL1* mutations were frameshift (n=6). Alterations in the *NPM1* gene were found in exons 6 and 12. One of them was homozygous and two were present in more than one individual. All the mutations observed in *NRAS* were single AA changes, with three being in the known codon 12 and only one in codon 146 (Table 13). These data are in line with the mutational profiles reported for acute myeloid leukemia and other myeloid neoplasms [39].

ZEB2 and *IKZF3*, transcription factors previously not reported as being mutated in human leukemias, were found to contain deleterious mutations in our BPDCN DNAs. *ZEB2* codes for a zinc finger/homeodomain protein that normally functions as a DNA-binding transcriptional repressor. We found two frameshift mutations and one single AA change in *ZEB2*). Finally, *IKZF3* and other members of the *IKAROS*-family of genes were also recurrently mutated in BPDCN. One patient had a frameshift mutation in *IKZF1*, another had a single AA change in *IKZF2* and three patients had alterations in *IKZF3*: one single AA change and one frameshift in two cases. Twenty percent of our BPDCN patients harbored mutations in *IKAROS*-family genes, suggesting a potential role of these genes in this specific type of myeloid leukemia.

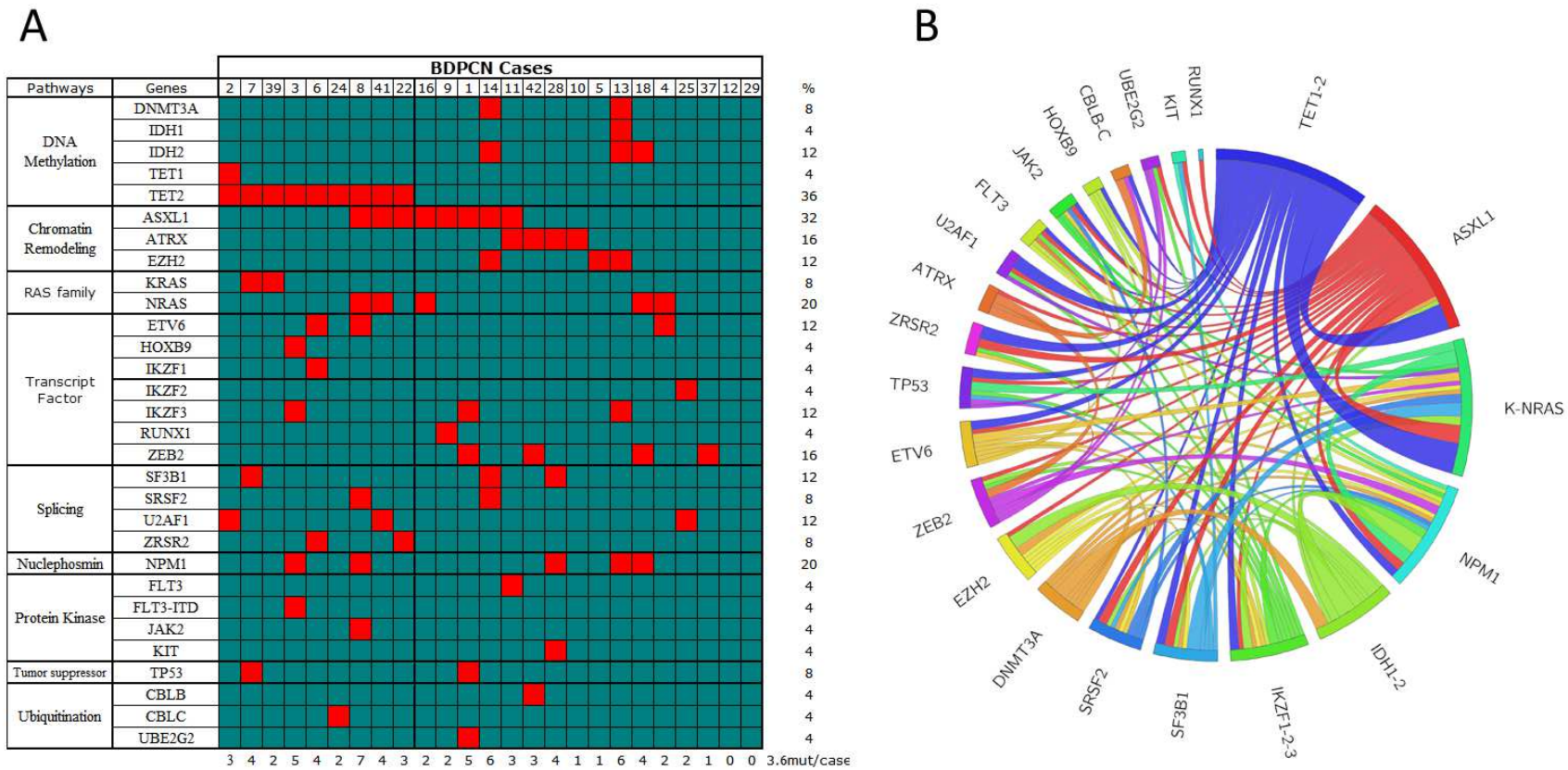

Figure 16: Recurrently mutated genes and their interaction in BPDCN. (A) 25 total cases of BPDCN were screened for somatic variants in 13 genes identified by WES and in 25 genes previously reported to be mutated in myeloid malignancies (a total of 38 genes). Mutations are represented in red. We observed mutations in 29 genes, among them, TET2 is the most frequently mutated (36%), followed by ASXL1 (32%), NRAS (20%) and NPM1 (20%), IKAROS family (20%) and ZEB2 (16%). (B) The circos diagram depicts the relative frequency and pairwise co-occurrence of mutations in the BPDCN patients. The length of the arc corresponds with the frequency of mutations in the first gene, and the width of the ribbon corresponds to the percentage of patients who also had a mutation in the second gene. Pairwise co-occurrence of mutations is denoted only once, beginning with the first gene in the clockwise direction. Mutations in TET1-2 and *IDH1-2* are mutually exclusive, as previously reported in other myeloid malignancies.

Because BPDCN is a rare disease that was recognized as a distinct entity only recently, information regarding genomic and genetic biomarkers is rather limited. Lucioni *et al*[14] studied 21 BPDCN patients by comparative genomic hybridization arrays and found complete or partial chromosomal losses with common deleted regions involving the *CDKN2A*/*CDKN2B*, *RB1*, *CDKN1B*), *LATS2* and *IKZF1* loci. Jardin *et al* [54] reported 54% and 38% of *TET2* and *TP53* mutations, respectively, in the 13 BPDCN individuals studied. In this study, we present, for the first time, a comprehensive analysis of BPDCN combining next generation sequencing with clinical characteristics and clinical outcomes of 25 patients. Although we did not find a common alteration that could explain the pathogenesis of BPDCN in the three patients studied by WES, we identified deleterious mutations in *IKZF3*, *HOXB9*, *UBE2G2* and *ZEB2* for the first time in human leukemia and more common alterations such as *TET1/2*, *ASXL1* and *U2AF1* previously reported as mutated in other myeloid neoplasms.

The best test of the relevance of individual mutations for pathogenesis (in the absence of functional validation) is their recurrence in other samples or other cancers. *TET2* mutations featured in our series appear similar to those observed in myeloid neoplasms: aberrations are mainly frameshift or nonsense and have been observed with the same frequencies (36-58%)[131]. However, our *TET2* and TP53 mutation frequency are lower compared to the BPDCN cases from Jardin *et al*. Similar to other hematological malignancies, the relevance of *TET2* mutations remains undetermined in this disease, which is characterized by a peculiar combination of various genomic alterations. Truncated exon 12 mutations in *ASXL1* have recently been described in 11% of patients with MDS, 43% of those with chronic myelomonocytic leukemia, 7% with primary and 47% with secondary AML [132]. The *NPM1* mutations was previouly shown to be present in 23.5% of AML patients and activating *RAS* mutations were found to be present in 10.7% [76].

Notably, only one BPDCN case displayed *JAK2* and other *FLT3*-ITD mutations, reinforcing that BPDCN and acute myeloid leukemia (AML) or myeloproliferative disorders may be genetically distinct. Pointing out to these differences, we found mutations in genes never shown to be involved in human leukemia. First, the *IKAROS* gene family encodes zinc-finger transcription factors that regulate gene expression via

chromatin remodeling in lymphoid development and differentiation. In fact, *IKZF1/2/3* mutations are prevalent in acute lymphoblastic leukemia (ALL) [132, 133]. Since 20% of our BPDCN patients harbored mutations in *IKAROS*-family genes, a potential role of these genes in normal or abnormal myelopoiesis should be considered. Secondly, although mutations in *ZEB2* have not been reported in human cancer, ZEB2 protein interacts with several members of the SMAD protein family that have been shown to play important roles in normal and malignant hematopoiesis. Finally, a preliminary report indicated that *ZEB2* was overexpressed in leukemia patients with *MLL* gene rearrangements [134].

Table 12: Pathogenic variants found using target next generation sequencing

Position	Gene	Ref	Variant	Var Freq	Coverage	Sample	Codon Change	AA Change	Type	Provean	SIFT
31022978	ASXL1	T	TA	45,57	1457	Case_09			Frameshift	NA	NA
31024844	ASXL1	G	GT	21,43	28	Case_11			Frameshift	NA	NA
31022898	ASXL1	TC	T	42,51	1583	Case_14			Frameshift	NA	NA
31025066	ASXL1	AG	A	21,43	392	Case_16			Frameshift	NA	NA
31021157	ASXL1	TTGTG	T	41,2	716	Case_22			Frameshift	NA	NA
31021162	ASXL1	T	A	42,9	613	Case_22	TTG TG[T/A] GTC	C387*	Nonsense	NA	NA
31023135	ASXL1	CCT	C	50,95	997	Case_41			Frameshift	NA	NA
31022550	ASXL1	G	T	49,66	149	Case_01	AGG [G/T]GA GGC	G679*	Nonsense	NA	NA
31024021	ASXL1	C	CT	27,19	1927	Case_08			Frameshift	NA	NA
76939409	ATRX	C	T	5,32	282	Case_10	GGA [G/A]AA AAA	E447K	Single AA Change	Neutral	Damaging
76939409	ATRX	C	T	5,71	210	Case_11	GGA [G/A]AA AAA	E447K	Single AA Change	Neutral	Damaging
76888705	ATRX	C	G	70,59	17	Case_28	GCT TT[G/C] GTT	L1708F	Single AA Change	Deleterious	Damaging
76776372	ATRX	G	A	5,49	1966	Case_42	CTC T[C/T]A GAG	S2365L	Single AA Change	Neutral	Damaging
76812991	ATRX	A	C	5,17	2497	Case_42	ACT TT[T/G] GAG	F2210L	Single AA Change	Deleterious	Tolerated
105377837	CBLB	C	CG	16,07	168	Case_42			Frameshift	NA	NA
45296756	CBLC	T	C	8,54	199	Case_24	CCC T[T/C]C TGC	F388S	Single AA Change	Deleterious	Damaging
25457243	DNMT3A	G	A	36,4	684	Case_13	AGC [C/T]GC TTG	R882C	Single AA Change	Deleterious	Damaging
25536822	DNMT3A	T	G	52,94	17	Case_14	GGG G[A/C]C ACC	D11A	Single AA Change	Neutral	Damaging
11992124	ETV6	TG	T	28,66	841	Case_06			Frameshift	NA	NA
11992124	ETV6	TG	T	18,47	1202	Case_04			Frameshift	NA	NA
11992124	ETV6	TG	T	26,2	1290	Case_08			Frameshift	NA	NA
148525964	EZH2	A	AAC	19,44	36	Case_13			Frameshift	NA	NA
148525964	EZH2	A	AAC	16,22	37	Case_14			Frameshift	NA	NA
148513776	EZH2	C	T	41,4	372	Case_05	CAC C[G/A]G TTG	R502Q	Single AA Change	Deleterious	Damaging
148523616	EZH2	G	C	29,92	1183	Case_05	TTA CA[C/G] TCC	H279Q	Single AA Change	Deleterious	Damaging
28602329	FLT3	G	A	12,33	981	Case_11	GGG G[C/T]G TGC	A680V	Single AA Change	Deleterious	Damaging
46700467	HOXB9	C	T	42,94	2718	Case_03	GCT C[G/A]C TCT	R183H	Single AA Change	Deleterious	Damaging
209113112	IDH1	C	T	19,42	515	Case_13	GGT C[G/A]T CAT	R132H	Single AA Change	Deleterious	Damaging
90631934	IDH2	C	T	14,87	2166	Case_13	ATC C[G/A]G AAC	R140Q	Single AA Change	Deleterious	Damaging
90631934	IDH2	C	T	45,51	2516	Case_14	ATC C[G/A]G AAC	R140Q	Single AA Change	Deleterious	Damaging
90631934	IDH2	C	T	5,54	1551	Case_18	ATC C[G/A]G AAC	R140Q	Single AA Change	Deleterious	Damaging
50450305	IKZF1	CA	C	54,59	850	Case_06			Frameshift	NA	NA
213921691	IKZF2	G	A	6,01	516	Case_25	GTG G[C/T]T GAC	A97V	Single AA Change	Neutral	Tolerated
37922598	IKZF3	C	CG	25,11	1561	Case_13			Frameshift	NA	NA
37922621	IKZF3	C	T	26,61	932	Case_01	GCC [G/A]AA GCC	E318K	Single AA Change	Deleterious	Damaging
37922598	IKZF3	C	CG	26,49	906	Case_03			Frameshift	NA	NA
5054619	JAK2	T	G	10,3	233	Case_08	ATT T[T/G]G ACA	L224W	Single AA Change	Deleterious	Damaging
55593690	KIT	A	T	4,64	3408	Case_28	CCC [A/T]GA AAC	R586*	Nonsense	NA	NA
25378562	KRAS	C	T	10,59	340	Case_07	TCA [G/A]CA AAG	A146T	Single AA Change	Deleterious	Damaging
25398282	KRAS	C	A	30,32	188	Case_07	GGT [G/T]GC GTA	G13C	Single AA Change	Deleterious	Damaging
25368410	KRAS	C	T	43,02	1311	Case_39	CCT [G/A]GC TGT	G179S	Single AA Change	Neutral	Damaging
170837544	NPM1	T	TCTGC	33,71	264	Case_13			Frameshift	NA	NA
170819932	NPM1	T	TTAG	100	30	Case_18	CTT [-/TAG] GCT	L158*	Nonsense	NA	NA
170819952	NPM1	AC	A	19,16	214	Case_18			Frameshift	NA	NA
170819930	NPM1	C	T	6,02	930	Case_28	AAA [C/T]TT GCT	L158F	Single AA Change	Neutral	Damaging
170837544	NPM1	T	TCTGC	33,83	133	Case_03			Frameshift	NA	NA
170819930	NPM1	C	T	5,83	103	Case_08	AAA [C/T]TT GCT	L158F	Single AA Change	Neutral	Damaging
115258747	NRAS	C	T	17,76	670	Case_16	GCA G[G/A]T GGT	G12R	Single AA Change	Deleterious	Damaging
115258748	NRAS	C	G	23,8	458	Case_18	GCA [G/C]GT GGT	G12R	Single AA Change	Deleterious	Damaging
115252204	NRAS	C	T	9,66	290	Case_41	TCA [G/A]CC AAG	A146T	Single AA Change	Deleterious	Damaging
115258747	NRAS	C	G	5,26	646	Case_04	GCA G[G/C]T GGT	G12A	Single AA Change	Deleterious	Damaging
115258747	NRAS	C	T	7,23	788	Case_08	GCA G[G/A]T GGT	G12D	Single AA Change	Deleterious	Damaging
36252864	RUNX1	T	TCGAC	45,29	223	Case_09			Frameshift	NA	NA
198266476	SF3B1	AT	A	19,52	210	Case_07			Frameshift	NA	NA
198266476	SF3B1	AT	A	16,32	380	Case_14			Frameshift	NA	NA
198257842	SF3B1	A	G	45,1	765	Case_28	GGT [T/C]GT GAA	C1204R	Single AA Change	Deleterious	Tolerated
74732959	SRSF2	G	T	45,29	170	Case_14	CGC C[C/A]C CCG	P122H	Single AA Change	Deleterious	Damaging
74732959	SRSF2	G	T	64,81	108	Case_08	CGC C[C/A]C CCG	P122H	Single AA Change	Deleterious	Damaging
70332921	TET1	C	T	8,95	525	Case_02	TCA [C/T]GA GTA	R276*	Nonsense	NA	NA
106157044	TET2	C	T	41,88	1466	Case_06	GAC [C/T]AA CAT	Q670*	Nonsense	NA	NA
106164901	TET2	A	ACGCTCACCAAT	36,9	626	Case_07			Frameshift	NA	NA
106197378	TET2	A	G	30	850	Case_07	CAG C[A/G]T AAG	H1925R	Single AA Change	Deleterious	Damaging
106180838	TET2	G	T	39,33	300	Case_22	GGT T[G/T]T TCA	C1310F	Single AA Change	Deleterious	Damaging
106196345	TET2	T	TA	37,59	689	Case_22			Frameshift	NA	NA
106156566	TET2	C	CT	29	1000	Case_24			Frameshift	NA	NA
106196415	TET2	C	A	39,23	808	Case_39	TCT T[C/A]A CAC	S1604*	Nonsense	NA	NA
106156566	TET2	C	CT	56,34	994	Case_41			Frameshift	NA	NA
106158403	TET2	AATAATTTT	A	36,54	572	Case_02			Frameshift	NA	NA
106164832	TET2	G	T	53,87	1253	Case_03	TGG [G/T]AA GGA	E1255*	Nonsense	NA	NA
106193892	TET2	C	T	27,58	765	Case_08	CGG [C/T]GA AAA	R1473*	Nonsense	NA	NA
7578536	TP53	T	G	38,65	727	Case_07	AAC [A/C]AG ATG	K132Q	Single AA Change	Deleterious	Damaging
7577551	TP53	C	T	62,63	835	Case_01	ATG [G/A]GC GGC	G244S	Single AA Change	Deleterious	Damaging
44514876	U2AF1	T	C	9,02	4091	Case_25	GCG G[A/G]A AAG	E124G	Single AA Change	Deleterious	Tolerated
44514769	U2AF1	T	TCTCATA	30,36	616	Case_41	GAG [-/TATGAG] ATG	E159EYE	Insertion	Deleterious	NA
44524456	U2AF1	G	A	25,27	182	Case_02	TGC T[C/T]T CGG	S34F	Single AA Change	Deleterious	Damaging
46197270	UBE2G2	T	A	28,68	129	Case_01	CTT G[A/T]T TAC	D63V	Single AA Change	Deleterious	Damaging
145156878	ZEB2	C	CG	27,98	193	Case_18			Frameshift	NA	NA
145156878	ZEB2	C	CG	23,91	46	Case_37			Frameshift	NA	NA
145161489	ZEB2	T	TC	18,2	1000	Case_42			Frameshift	NA	NA
145157495	ZEB2	A	C	36,05	491	Case_01	GGG C[T/G]T GGA	L420R	Single AA Change	Neutral	Damaging
15827389	ZRSR2	C	T	59,74	77	Case_06	GAT [C/T]GA GCT	R169*	Nonsense	NA	NA
15834002	ZRSR2	A	AG	17,42	666	Case_22			Frameshift	NA	NA

Table 13: The mutated genes in BPDCN

Annotated gene	Mutation type	Position	Allele change	AA Change	Case numbers with mutation	Frequency (%)
TET2	Nonsense	106157044	C>T	p.Gln649*	1 / 25	9 / 25 (36%)
	Frameshift	106164901	A>ACGCTCACCAAT	p.Arg1262Serfs*8	1 / 25	
	Single AA Change	106197378	A>G	p.His1904Arg	1 / 25	
	Single AA Change	106180838	G>T	p.Cys1289Phe	1 / 25	
	Frameshift	106196345	T>TA	p.Tyr1560*	1 / 25	
	Frameshift	106156566	C>CT	p.Ile490Tyrfs*14	2 / 25	
	Nonsense	106196415	C>A	p.Ser1583*	1 / 25	
	Frameshift	106158403	AATAATTTT>A	p.Asn1103Argfs*24	1 / 25	
	Nonsense	106164832	G>T	p.Glu1234*	1 / 25	
	Nonsense	106193892	C>T	p.Arg1452*	1 / 25	
ASXL1	Frameshift	31022978	T>TA	p.Thr822Asnfs*11	1 / 25	8 / 25 (32%)
	Frameshift	31024844	G>GT	p.Leu1444Serfs*3	1 / 25	
	Frameshift	31022898	TC>T	p.Trp796Glyfs*22	1 / 25	
	Frameshift	31025066	AG>A	p.Gly1518Alafs*23	1 / 25	
	Frameshift	31021157	TTGTG>T	p.Cys387Serfs*74	1 / 25	
	Nonsense	31021162	T>A	p.Cys387*	1 / 25	
	Frameshift	31023135	CCT>C	p.Pro874Hisfs*5	1 / 25	
	Nonsense	31022550	G>T	p.Gly679*	1 / 25	
	Frameshift	31024021	C>CT	p.Leu1170Phefs*12	1 / 25	
NPM1	Frameshift	170837544	T>TCTGC	p.Trp288Cysfs*12	2 / 25	5 / 25 (20%)
	Nonsense	170819932	T>TTAG	p.Ala159Argfs*35	1 / 25	
	Frameshift	170819952	AC>A	p.Asp165Glufs*28	1 / 25	
	Single AA Change	170819930	C>T	p.Leu158Phe	2 / 25	
NRAS	Single AA Change	115258747	C>T	p.Gly12Asp	2 / 25	5 / 25 (20%)
	Single AA Change	115258748	C>G	p.Gly12Arg	1 / 25	
	Single AA Change	115252204	C>T	p.Ala146Thr	1 / 25	
	Single AA Change	115258747	C>G	p.Gly12Asp	1 / 25	
IKZF1	Frameshift	50450305	CA>C	p.Ile164Serfs*29	1 / 25	5 / 25 (20%)
IKZF2	Single AA Change	213921691	G>A	p.Ala91Val	1 / 25	
IKZF3	Single AA Change	37922621	C>T	p.Glu318Lys	1 / 25	
	Frameshift	37922598	C>CG	p.Thr326Hisfs*25	2 / 25	
ZEB2	Single AA Change	145157495	A>C	p.Leu420Arg	1 / 25	4 / 25 (16%)
	Frameshift	145156878	C>CG	p.Gly626Arg	2 / 25	
	Frameshift	145161489	T>TC	p.Asp268Argfs*12	1 / 25	
HOXB9	Single AA Change	46700467	C>T	p.Arg183His	1 / 25	1 / 25 (4%)
UBE2G2	Single AA Change	46197270	T>A	p.Asp63Val	1 / 25	1 / 25 (4%)

4.3.3 Functional categorization of mutated genes

We used circus diagrams to assess patterns of mutual exclusivity and co-occurrence between sets of genes. We identified mutually exclusive alterations that affect *TET1/2* and *IDH1/2*; *TET1/2* and *DNMT3A;* and between *IKAROS* family genes. Additionally, mutations in *ASXL1* and *DNMT3A* rarely co-occurred in our samples, with only one case harboring mutations in both. A rare co-occurrence was also observed between the splicing genes with only one case having mutations in both *SF3B1* and *SRSF2* (Figure 16B and Figure 17).

We also grouped mutations into larger sets or pathways and examined patterns of mutual exclusivity and co-occurrence between these groups (Figure 16A and Figure 18). Out of the 25 samples, 23 (92%) contained at least one mutation in one of the

nine categories. The nine categories and how frequently they were mutated are: aberrant methylation (48%), chromatin remodeling (52%), transcription factors (44%), splicing (36%), protein kinase (16%), ubiquitination (12%), *NPM1* (20%), *RAS* (28%) and *TP53* (8%). Half of the tumors had mutations affecting the methylation profile and chromatin remodeling pathways, and 20% had mutations in genes in both pathways.

Recently, a series of mutually exclusive mutations that affect *TET2*, *IDH1* or *IDH2* mutations, and potentially other as yet unidentified disease alleles, was reported [39, 135]. We observed the same pattern of mutual exclusiveness in our series and additionally between *IKAROS* genes.

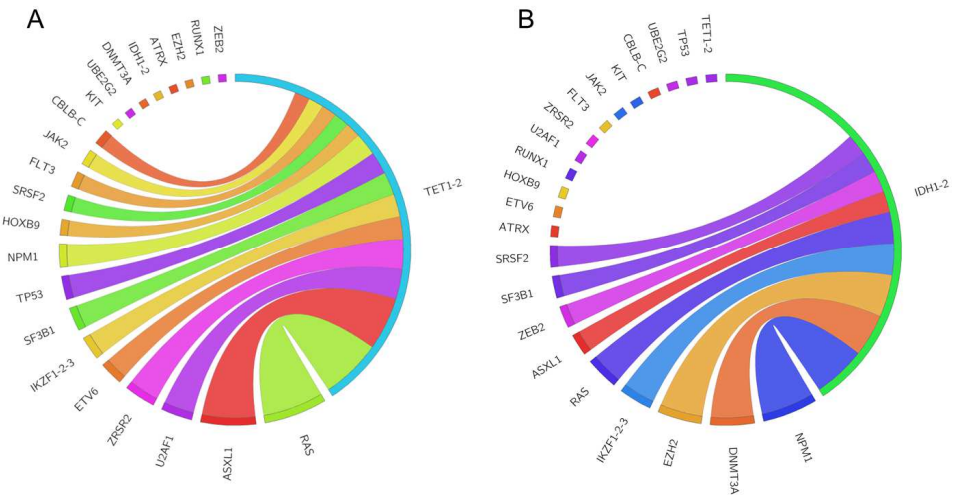

Figure 17: Pairwise co-occurrence of (A) *TET1-2* and (B) *IDH1-2* mutations in BPDCN. A circos diagram depicts the relative frequency and pairwise co-occurrence of other mutations with *TET1-2 or IDH1-2* mutations in the entire cohort. The length of the arc indicates the frequency of mutations in the first gene, and the width of the ribbon represents the percentage of BPDCN patients with *TET1-2 or IDH1-2* mutations who bear the second gene mutation. Pairwise co-occurrence of mutations is denoted only once, beginning with the *TET1-2 or IDH1-2* gene in the clockwise direction.

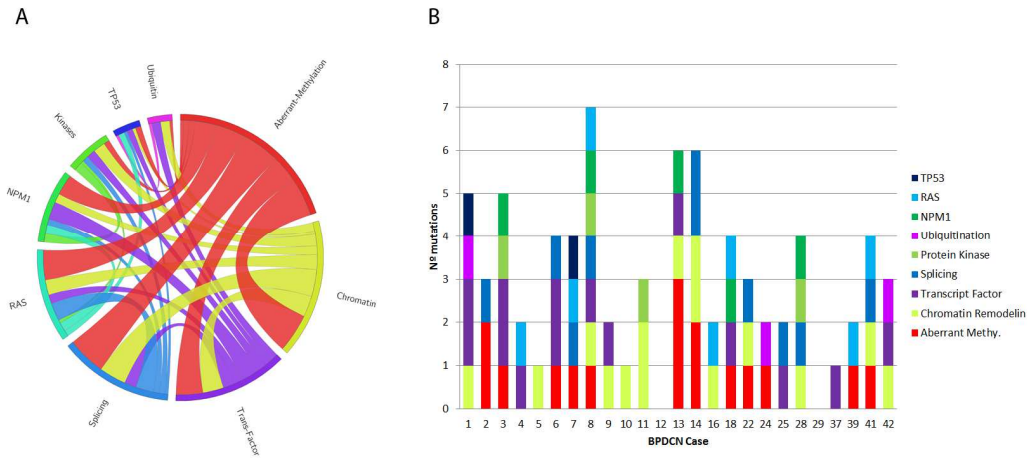

Figure 18: Target NGS result by pathways. (A) A Circos diagram depicts the relative frequency and pairwise co-occurrence of mutations by pathways in the BPDCN patients. The length of the arc corresponds to the frequency of mutations in the first pathway, and the width of the ribbon corresponds to the percentage of patients who also had a mutation in the second pathway. Pairwise co-occurrence of mutations is denoted only once, beginning with the first pathway in the clockwise direction; (B) Distribution of numbers and categories of unknown and pathogenic mutations among the 25 BPDCN cases. We do not identify any unknown nor pathogenic mutations among case 12 and 29 in the 38 studied genes, however, in case 8 we found mutations in almost all the represented pathways.

4.3.4 Clinical relevance of aberrant methylation genes mutations

The analysis of survival data from the study cohort revealed several significant differences among the patients. First, we found that the patients with mutations in genes in methylation pathways had a significantly reduced OS rate compared to individuals without mutations in these genes (median 11 months versus 79 months; p=0.047, Figure 19A). On the other hand, patients with mutations in *ETV6* (p=0.022), *TP53* (p=0.003), and N/H*RAS* (p=0.015) had either a statistically significant worse prognosis (Figure 20) or showed a trend in the same direction (p=0.073) when a mutation of *IKZF1/2/3* was present. Taking into consideration the limited sample size, we also combined the subgroups based on the functional classes for the outcome analysis. Among the combinations with clinical interest, the worst clinically significant changes were those affecting the RAS family, TP53 and other transcription factors (median 15 months, versus 99 months for the non-mutated cases; p=0.013) (Figure 19B). These data suggest that mutations in aberrant methylation genes represent a poor prognostic factor and are probably involved in the pathogenesis of BPDCN.

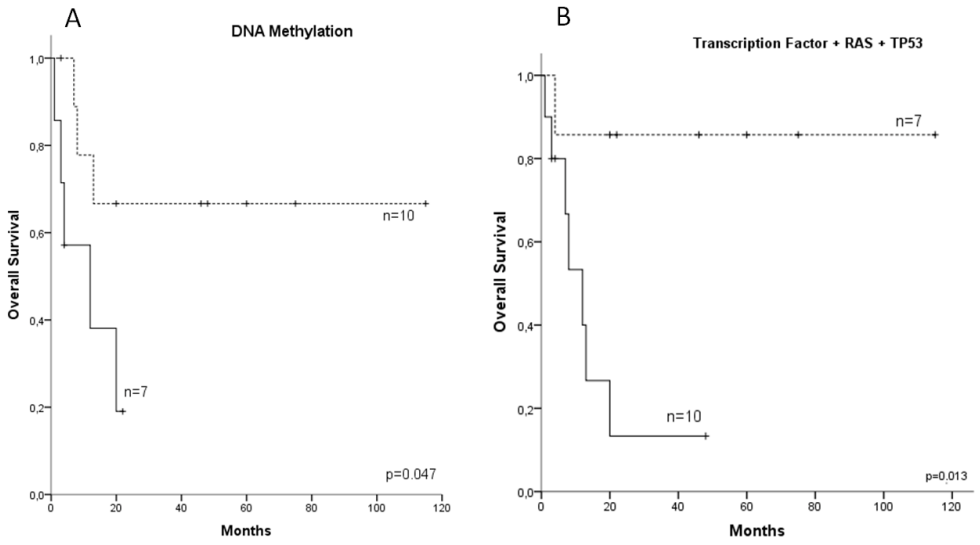

Figure 19: Kaplan-Meier overall survival (OS) curves. (A) We observed a significant difference in the OS according to the mutational status of genes included in the DNA methylation class (P=0.047; mut: n=7 and 11 months; wt: n=10 and 79 months); (B) We observed a significant difference in the OS according to the mutational status of genes included in the transcription factor class / *TP53* / *RAS* (P=0.013; mut: n=10 and 15 months; wt: n=7 and 99 months).

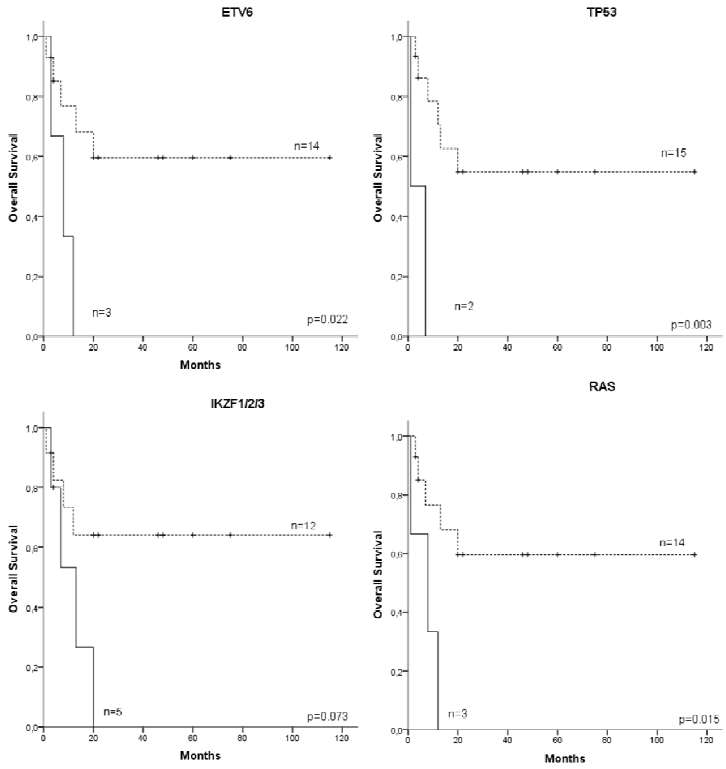

Figure 20: Kaplan-Meier overall survival curves for *ETV6* (p=0,022), *TP53* (p=0,003), *IKZF1/2/3* (p=0,073) and *RAS* (p=0,015). Patients with mutations in these genes had a reduced cumulative survival. As the number in each subgroup of patients with individual gene mutations was very small, we combined the subgroups based on the functional classes for the outcome analysis.

Correlation between epigenetic modifiers mutations and prognosis has been controversial. *TET2* mutations are associated with poor prognosis in normal karyotype-AML [136], but with no clear association in myelodisplastic syndromes (MDS) [137] or myeloproliferative neoplasm (MPN) [138]. Alterations in *IDH1/2* are associated with a favorable or negligible effect on prognosis, depending on the type of mutation in AML [139, 140]. No importance was observed in MDS [137] or MPN [138, 141]. Several studies show that *DNMT3A* mutations are associated with worsened overall survival in patients with AML [70, 79] and an adverse prognosis in MDS [142]. In our series of BPDCN patients, we found that patients with mutations in genes in methylation pathways had a significantly reduced overall survival as compared with individuals without mutations in these genes. These data suggest that mutations in DNA methylation genes represent a poor prognostic factor but are likely to be involved in the pathogenesis of BPDCN.

Changing the paradigm of what is known about BPDCN, our data suggest that this dendritic cell leukemia has a mutational profile strikingly similar to that of other

well-defined myeloid leukemias. Despite the numerous advances in our understanding of the genetics and epigenetics of myeloid malignancies, only limited examples of clinical translation of these findings into new therapeutic approaches have been implemented. However, it has now become clear that epigenetic modifiers provide new targets for therapeutic intervention and that targeting these enzymatic activities are currently being explored from a therapeutic standpoint in several types of leukemia[129, 130]. Our study provides, for the first time, molecular data that suggest that a complete and innovative change is required in the treatment approach adopted for patients with BPDCN, including the administration of drugs that target aberrant methylation, chromatin remodeling, specific transcription factors and aberrant splicing.

CHAPTER 5: Conclusions
Conclusiones

- We have been able to successfully apply next generation technologies to achieve deeper insights into the genetic pathogenesis of rare myeloid disorders and to integrate the obtained data in order to generate new and better genetic biomarkers that can be used in the clinical management these patients. In detail:
- Whole-exome sequencing allowed the identification of a large number of mutated genes, some of them with prognostic and predictive significance, such as *ASXL1* and *TP53* in both chronic phase and blastic crisis of chronic myeloid leukemia. The study of the mutation profile through the course of the disease indicated that, at least in this patient, the number and the type of mutations were similar at chronic phase and blastic crisis. In addition, we identified for the first time deleterious mutations in *IKZF3*, *UBE2G2* and *ZEB2*.
- While current diagnostic procedures recommend the study of *ABL1* mutations in non-responders chronic myeloid leukemia patients, our data suggest that sequencing a wider panel of genes, that includes *ASXL1*, *TP53* and *IKZF* gene family, could be also beneficial in the clinical management of these patients.
- In addition to the recurrent *CSF3R* mutation, we found and validated mutations in *U2AF1*, *TET2*, *LUC7L2* and *ASXL1* in a chronic neutrophilic patient. We also observed a large genomic segment displaying uniparental disomy in chromosome 7q, including the locus of *LUC7L2,* demonstrating a new mechanism of pathogenesis for this gene. Our study provides, for the first time, a massive molecular and expression data, revealing a large amount of genomic alterations in chronic neutrophilic leukemia.
- Functional classification of genes differently expressed between chronic neutrophilic leukemia cells and normal control revealed an enrichment of categories like cell signaling; cell death and survival; and gene expression.
- Although no clear genome-wide increase in intron retention was observed in the chronic neutrophilic leukemia cells, we found an altered pattern of splicing in *RUNX1* gene. In addition, we identified an expressed fusion gene *PIM3-SCO2* that resulted from an inversion in chromosome 22. In conclusion, we show the concurrence of several genetic mechanisms of mutation that cooperate with the *CSF3R* mutation in chronic neutrophilic leukemia. In this complex scenario,

a combination of new targets therapies may be considered as reasonable option for the therapeutic management of this aggressive and rare subtype of leukemia.

- Whole-exome sequencing of three patients with blastic plasmacytoid dendritic cell neoplasm revealed no common affected genes between patients, but a clear overlap in terms of molecular and disease pathways. Changing the paradigm of what is known about this disease, our data suggest that this dendritic cell leukemia has a mutational profile strikingly similar to that of other well-defined myeloid leukemias.
- We successfully design and applied a comprehensive and cost effective targeted deep re-sequencing for screening 25 blastic plasmacytoid dendritic cell neoplasm tumors. Recurrence of many genes was found and ranged in prevalence from 36% for previously known genes, such as *TET2,* to 12-16% for newly identified genes, such as *IKZF3 or ZEB2.* Half of the tumors had mutations affecting the methylation profile and the chromatin remodeling pathways.
- The analysis of clinical data from the blastic plasmacytoid dendritic cell neoplasm cohort revealed that patients with mutations in genes in the methylation pathways had a significantly reduced cumulative survival. Our study provides, for the first time, molecular data that suggest that a complete and innovative change is required in the treatment approach adopted for patients with blastic plasmacytoid dendritic cell neoplasm, including the administration of drugs that target aberrant methylation, chromatin remodeling, specific transcription factors and aberrant splicing.

- Hemos sido capaces de aplicar con éxito tecnologías de secuenciación completa de última generación para lograr una visión más profunda de la patogénesis genética de enfermedades mieloides raras. Además fuimos capaces de integrar los datos obtenidos con la finalidad de generar nuevos y mejores biomarcadores genéticos que se pueden utilizar en el manejo clínico de estos pacientes. En detalle:
- La secuenciación del exoma completo permitió la identificación de un gran número de genes mutados, algunos de ellos con significado pronóstico y predictivo, tales como *ASXL1* y *TP53*, tanto en la fase crónica como en la crisis blástica de la leucémia mieloide crónica. El estudio del perfil de mutacional a lo largo de la progresión de la enfermedad indica que, al menos en este paciente, el número y el tipo de mutaciones fueron similares en ambas fases. Además, se identificaron por primera vez mutaciones deletéreas en *IKZF3*, *UBE2G2* y *ZEB2*.
- Aunque los procedimientos de diagnóstico actuales recomiendan el estudio de mutaciones de *ABL1* en pacientes con leucemia mieloide crónica no respondedores a los inibidores de las tirosina quinasas, nuestros datos sugieren que la secuenciación de un panel más amplio de genes, que incluyan *ASXL1*, *TP53* y la familia de genes *IKZF*, podría ser también beneficiosa en el manejo clínico de estos pacientes.
- Además de la mutación recurrente en *CSF3R*, encontramos y validamos mutaciones en *U2AF1*, *TET2*, *LUC7L2* y *ASXL1*, en un paciente con leucémia neutrofílica crónica. También se observó una disomía uniparental en el brazo largo del cromosoma 7, incluyendo el locus de *LUC7L2*, lo que señala un nuevo mecanismo por la patogénesis de este gen. Nuestro estudio proporciona, por primera vez, un conjunto de datos moleculares y de expresión masivos, que revela una gran cantidad de alteraciones genómicas en la leucemia neutrofílica crónica.
- La clasificación funcional de los genes diferencialmente expresados entre las células de leucemia neutrofílica crónica y las del control normal reveló un enriquecimiento de categorías como la señalización celular, la muerte y supervivencia celular.

- Aunque no se observó un claro aumento en la inestabilidad global de los proceso de *splicing* en las células de la leucemia neutrofílica crónica, encontramos un patrón alterado de empalme en el gen *RUNX1*. Además, hemos identificado la expresión de un gen de fusión, *PIM3-SCO2*, que resultó de una inversión en el brazo largo del cromosoma 22. En conclusión, se demuestra la concurrencia de varios mecanismos genéticos que cooperan con la mutación en *CSF3R* en la leucemia neutrofílica crónica. En este complejo ensenario, una combinación de nuevas terapias específicas puede ser considerada como opción razonable para el manejo terapéutico de este subtipo agresivo y raro de leucemia.
- La secuenciación del exoma de tres pacientes con neoplasia blástica de células dendríticas no reveló genes afectados comunes entre los pacientes, sino una clara superposición en términos de vías moleculares. Cambiando el paradigma de lo que se conoce acerca de esta enfermedad, nuestros datos sugieren que este tipo de leucemia de células dendríticas tiene un perfil mutacional sorprendentemente similar a la de otras leucemias mieloides bien definidas.
- Diseñamos y aplicamos con éxito una re-secuenciación dirigida de un amplio panel de genes para la investigación de 25 tumores de neoplasias de células dendríticas. La recurrencia de muchos genes varió en la prevalencia del 36% para los genes previamente conocidos, como *TET2*, y el 12-16% para los genes recientemente identificados, tales como *IKZF3* o *ZEB2*. La mitad de los tumores tenían mutaciones que afectan el perfil de metilación y las vías de remodelación de la cromatina.
- El análisis de los datos clínicos de la cohorte de neoplasia de células dendríticas reveló que los pacientes con mutaciones en los genes en las vías de metilación tenían una supervivencia acumulada significativamente reducida. Nuestro estudio proporciona, por primera vez, los datos moleculares que sugieren que se requiere un cambio completo e innovador en el enfoque de tratamiento adoptado para los pacientes con neoplasia de células dendríticas, incluyendo la administración de fármacos que se dirigen contra la metilación aberrante y remodelación de la cromatina.

CHAPTER 6: References

1. Tefferi, A. and J.W. Vardiman, *Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms.* Leukemia, 2008. **22**(1): p. 14-22.
2. Cools, J., et al., *A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome.* N Engl J Med, 2003. **348**(13): p. 1201-14.
3. James, C., et al., *A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera.* Nature, 2005. **434**(7037): p. 1144-8.
4. Kralovics, R., et al., *A gain-of-function mutation of JAK2 in myeloproliferative disorders.* N Engl J Med, 2005. **352**(17): p. 1779-90.
5. Tefferi, A., T.L. Lasho, and G. Gilliland, *JAK2 mutations in myeloproliferative disorders.* N Engl J Med, 2005. **353**(13): p. 1416-7; author reply 1416-7.
6. Levine, R.L., et al., *Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis.* Cancer Cell, 2005. **7**(4): p. 387-97.
7. Pardanani, A.D., et al., *MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients.* Blood, 2006. **108**(10): p. 3472-6.
8. Pardanani, A., C. Akin, and P. Valent, *Pathogenesis, clinical features, and treatment advances in mastocytosis.* Best Pract Res Clin Haematol, 2006. **19**(3): p. 595-615.
9. Loh, M.L., et al., *Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis.* Blood, 2004. **103**(6): p. 2325-31.
10. Stephens, K., et al., *Interstitial uniparental isodisomy at clustered breakpoint intervals is a frequent mechanism of NF1 inactivation in myeloid malignancies.* Blood, 2006. **108**(5): p. 1684-9.
11. Tartaglia, M., et al., *Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia.* Nat Genet, 2003. **34**(2): p. 148-50.
12. Willman, C.L., *Molecular genetic features of myelodysplastic syndromes (MDS).* Leukemia, 1998. **12 Suppl 1**: p. S2-6.
13. Hirsch-Ginsberg, C., et al., *RAS mutations are rare events in Philadelphia chromosome-negative/bcr gene rearrangement-negative chronic myelogenous leukemia, but are prevalent in chronic myelomonocytic leukemia.* Blood, 1990. **76**(6): p. 1214-9.
14. Padua, R.A., et al., *RAS mutations in myelodysplasia detected by amplification, oligonucleotide hybridization, and transformation.* Leukemia, 1988. **2**(8): p. 503-10.
15. Mrozek, K. and C.D. Bloomfield, *Chromosome aberrations, gene mutations and expression changes, and prognosis in adult acute myeloid leukemia.* Hematology Am Soc Hematol Educ Program, 2006: p. 169-77.
16. Mrozek, K., et al., *Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification?* Blood, 2007. **109**(2): p. 431-48.
17. Paschka, P., et al., *Adverse prognostic significance of KIT mutations in adult acute myeloid leukemia with inv(16) and t(8;21): a Cancer and Leukemia Group B Study.* J Clin Oncol, 2006. **24**(24): p. 3904-11.

18. Falini, B., et al., *Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype.* N Engl J Med, 2005. **352**(3): p. 254-66.
19. Dohner, K., et al., *Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations.* Blood, 2005. **106**(12): p. 3740-6.
20. Greene, M.E., et al., *Mutations in GATA1 in both transient myeloproliferative disorder and acute megakaryoblastic leukemia of Down syndrome.* Blood Cells Mol Dis, 2003. **31**(3): p. 351-6.
21. Vardiman, J.W., et al., *The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes.* Blood, 2009. **114**(5): p. 937-51.
22. Sawyers, C.L., *Chronic myeloid leukemia.* N Engl J Med, 1999. **340**(17): p. 1330-40.
23. Hochhaus, A., et al., *Six-year follow-up of patients receiving imatinib for the first-line treatment of chronic myeloid leukemia.* Leukemia, 2009. **23**(6): p. 1054-61.
24. Druker, B.J., et al., *Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia.* N Engl J Med, 2006. **355**(23): p. 2408-17.
25. Barnes, D.J. and J.V. Melo, *Primitive, quiescent and difficult to kill: the role of non-proliferating stem cells in chronic myeloid leukemia.* Cell Cycle, 2006. **5**(24): p. 2862-6.
26. Apperley, J.F., *Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia.* Lancet Oncol, 2007. **8**(11): p. 1018-29.
27. Reilly, J.T., *Chronic neutrophilic leukaemia: a distinct clinical entity?* Br J Haematol, 2002. **116**(1): p. 10-8.
28. Bohm, J. and H.E. Schaefer, *Chronic neutrophilic leukaemia: 14 new cases of an uncommon myeloproliferative disease.* J Clin Pathol, 2002. **55**(11): p. 862-4.
29. Elliott, M.A., et al., *WHO-defined chronic neutrophilic leukemia: a long-term analysis of 12 cases and a critical review of the literature.* Leukemia, 2005. **19**(2): p. 313-7.
30. Breccia, M., et al., *Identification of risk factors in atypical chronic myeloid leukemia.* Haematologica, 2006. **91**(11): p. 1566-8.
31. Kurzrock, R., et al., *BCR rearrangement-negative chronic myelogenous leukemia revisited.* J Clin Oncol, 2001. **19**(11): p. 2915-26.
32. Martiat, P., J.L. Michaux, and J. Rodhain, *Philadelphia-negative (Ph-) chronic myeloid leukemia (CML): comparison with Ph+ CML and chronic myelomonocytic leukemia. The Groupe Francais de Cytogenetique Hematologique.* Blood, 1991. **78**(1): p. 205-11.
33. Hernandez, J.M., et al., *Clinical, hematological and cytogenetic characteristics of atypical chronic myeloid leukemia.* Ann Oncol, 2000. **11**(4): p. 441-4.
34. Zhang, X., J. Pan, and J. Guo, *Presence of the JAK2 V617F Mutation in a Patient with Chronic Neutrophilic Leukemia and Effective Response to Interferon Alfa-2b.* Acta Haematol. **130**(1): p. 44-6.
35. Bohm, J., et al., *Evidence of clonality in chronic neutrophilic leukaemia.* J Clin Pathol, 2003. **56**(4): p. 292-5.
36. Kwong, Y.L. and G. Cheng, *Clonal nature of chronic neutrophilic leukemia.* Blood, 1993. **82**(3): p. 1035-6.

37. Pane, F., et al., *Neutrophilic-chronic myeloid leukemia: a distinct disease with a specific molecular marker (BCR/ABL with C3/A2 junction).* Blood, 1996. **88**(7): p. 2410-4.
38. Maxson, J.E., et al., *Oncogenic CSF3R mutations in chronic neutrophilic leukemia and atypical CML.* N Engl J Med. **368**(19): p. 1781-90.
39. Steensma, D.P., et al., *The JAK2 V617F activating tyrosine kinase mutation is an infrequent event in both "atypical" myeloproliferative disorders and myelodysplastic syndromes.* Blood, 2005. **106**(4): p. 1207-9.
40. Vardiman, J. and E. Hyjek, *World health organization classification, evaluation, and genetics of the myeloproliferative neoplasm variants.* Hematology Am Soc Hematol Educ Program, 2011. **2011**: p. 250-6.
41. Herling, M. and D. Jones, *CD4+/CD56+ hematodermic tumor: the features of an evolving entity and its relationship to dendritic cells.* Am J Clin Pathol, 2007. **127**(5): p. 687-700.
42. Willemze, R., et al., *WHO-EORTC classification for cutaneous lymphomas.* Blood, 2005. **105**(10): p. 3768-85.
43. Cella, M., et al., *Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon.* Nat Med, 1999. **5**(8): p. 919-23.
44. Chaperot, L., et al., *Identification of a leukemic counterpart of the plasmacytoid dendritic cells.* Blood, 2001. **97**(10): p. 3210-7.
45. Petrella, T., et al., *'Agranular CD4+ CD56+ hematodermic neoplasm' (blastic NK-cell lymphoma) originates from a population of CD56+ precursor cells related to plasmacytoid monocytes.* Am J Surg Pathol, 2002. **26**(7): p. 852-62.
46. Pilichowska, M.E., et al., *CD4+/CD56+ hematodermic neoplasm ("blastic natural killer cell lymphoma"): neoplastic cells express the immature dendritic cell marker BDCA-2 and produce interferon.* Am J Clin Pathol, 2007. **128**(3): p. 445-53.
47. Feuillard, J., et al., *Clinical and biologic features of CD4(+)CD56(+) malignancies.* Blood, 2002. **99**(5): p. 1556-63.
48. Petrella, T., et al., *Blastic NK-cell lymphomas (agranular CD4+CD56+ hematodermic neoplasms): a review.* Am J Clin Pathol, 2005. **123**(5): p. 662-75.
49. Reichard, K.K., et al., *CD4(+) CD56(+) lineage-negative malignancies are rare tumors of plasmacytoid dendritic cells.* Am J Surg Pathol, 2005. **29**(10): p. 1274-83.
50. Lee, J.K. and G. Schiller, *Blastic plasmacytoid dendritic cell neoplasm.* Clin Adv Hematol Oncol. **10**(1): p. 60-2.
51. Marafioti, T., et al., *Novel markers of normal and neoplastic human plasmacytoid dendritic cells.* Blood, 2008. **111**(7): p. 3778-92.
52. Montes-Moreno, S., et al., *SPIB, a novel immunohistochemical marker for human blastic plasmacytoid dendritic cell neoplasms: characterization of its expression in major hematolymphoid neoplasms.* Blood. **121**(4): p. 643-7.
53. Lucioni, M., et al., *Twenty-one cases of blastic plasmacytoid dendritic cell neoplasm: focus on biallelic locus 9p21.3 deletion.* Blood. **118**(17): p. 4591-4.
54. Jardin, F., et al., *TET2 and TP53 mutations are frequently observed in blastic plasmacytoid dendritic cell neoplasm.* Br J Haematol. **153**(3): p. 413-6.
55. Reimer, P., et al., *What is CD4+CD56+ malignancy and how should it be treated?* Bone Marrow Transplant, 2003. **32**(7): p. 637-46.

56. Assaf, C., et al., *CD56-positive haematological neoplasms of the skin: a multicentre study of the Cutaneous Lymphoma Project Group of the European Organisation for Research and Treatment of Cancer.* J Clin Pathol, 2007. **60**(9): p. 981-9.
57. Dalle, S., et al., *Blastic plasmacytoid dendritic cell neoplasm: is transplantation the treatment of choice?* Br J Dermatol. **162**(1): p. 74-9.
58. Roos-Weil, D., et al., *Stem cell transplantation can provide durable disease control in blastic plasmacytoid dendritic cell neoplasm: a retrospective study from the European Group for Blood and Marrow Transplantation.* Blood. **121**(3): p. 440-6.
59. Nowell, P.C., *The minute chromosome (Phl) in chronic granulocytic leukemia.* Blut, 1962. **8**: p. 65-6.
60. Godley, L.A., et al., *An integrated genomic approach to the assessment and treatment of acute myeloid leukemia.* Semin Oncol. **38**(2): p. 215-24.
61. Suela, J., S. Alvarez, and J.C. Cigudosa, *DNA profiling by arrayCGH in acute myeloid leukemia and myelodysplastic syndromes.* Cytogenet Genome Res, 2007. **118**(2-4): p. 304-9.
62. Watt, C.D. and A. Bagg, *Molecular diagnosis of acute myeloid leukemia.* Expert Rev Mol Diagn. **10**(8): p. 993-1012.
63. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
64. Meyerson, M., S. Gabriel, and G. Getz, *Advances in understanding cancer genomes through second-generation sequencing.* Nat Rev Genet. **11**(10): p. 685-96.
65. Metzker, M.L., *Sequencing technologies - the next generation.* Nat Rev Genet. **11**(1): p. 31-46.
66. Blencowe, B.J., S. Ahmad, and L.J. Lee, *Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes.* Genes Dev, 2009. **23**(12): p. 1379-86.
67. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences.* Nature, 2009. **462**(7271): p. 315-22.
68. Neff, T. and S.A. Armstrong, *Chromatin maps, histone modifications and leukemia.* Leukemia, 2009. **23**(7): p. 1243-51.
69. Ley, T.J., et al., *DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome.* Nature, 2008. **456**(7218): p. 66-72.
70. Ley, T.J., et al., *DNMT3A mutations in acute myeloid leukemia.* N Engl J Med. **363**(25): p. 2424-33.
71. Fong, P.C., et al., *Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers.* N Engl J Med, 2009. **361**(2): p. 123-34.
72. Iyer, L., et al., *Genetic predisposition to the metabolism of irinotecan (CPT-11). Role of uridine diphosphate glucuronosyltransferase isoform 1A1 in the glucuronidation of its active metabolite (SN-38) in human liver microsomes.* J Clin Invest, 1998. **101**(4): p. 847-54.
73. Weinshilboum, R.M. and L. Wang, *Pharmacogenetics and pharmacogenomics: development, science, and translation.* Annu Rev Genomics Hum Genet, 2006. **7**: p. 223-45.

74. Birch, J.M., et al., *Relative frequency and morphology of cancers in carriers of germline TP53 mutations.* Oncogene, 2001. **20**(34): p. 4621-8.
75. Song, W.J., et al., *Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia.* Nat Genet, 1999. **23**(2): p. 166-75.
76. Mardis, E.R., et al., *Recurring mutations found by sequencing an acute myeloid leukemia genome.* N Engl J Med, 2009. **361**(11): p. 1058-66.
77. Welch, J.S., et al., *Use of whole-genome sequencing to diagnose a cryptic fusion oncogene.* JAMA. **305**(15): p. 1577-84.
78. Link, D.C., et al., *Identification of a novel TP53 cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML.* JAMA. **305**(15): p. 1568-76.
79. Yan, X.J., et al., *Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia.* Nat Genet. **43**(4): p. 309-15.
80. Damm, F., et al., *Integrative prognostic risk score in acute myeloid leukemia with normal karyotype.* Blood. **117**(17): p. 4561-8.
81. Schnittger, S., et al., *IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutated NPM1 status.* Blood. **116**(25): p. 5486-96.
82. Figueroa, M.E., et al., *Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation.* Cancer Cell. **18**(6): p. 553-67.
83. Ko, M., et al., *Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2.* Nature. **468**(7325): p. 839-43.
84. Yamashita, Y., et al., *Array-based genomic resequencing of human leukemia.* Oncogene. **29**(25): p. 3723-31.
85. *Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia.* N Engl J Med, 2013. **368**(22): p. 2059-74.
86. Bejar, R., et al., *Clinical effect of point mutations in myelodysplastic syndromes.* N Engl J Med, 2011. **364**(26): p. 2496-506.
87. Offit, K., et al., *Cytogenetic analysis of 434 consecutively ascertained specimens of non-Hodgkin's lymphoma: clinical correlations.* Blood, 1991. **77**(7): p. 1508-15.
88. Andrews, S. *FastQC, a quality control tool for high throughput sequence data*. 2012; Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
89. Rubio-Camarillo, M., et al., *RUbioSeq: a suite of parallelized pipelines to automate exome variation and bisulfite-seq analyses.* Bioinformatics.
90. Jones, S., et al., *Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma.* Science. **330**(6001): p. 228-31.
91. Fernandez-Mercado, M., et al., *Mutation patterns of 16 genes in primary and secondary acute myeloid leukemia (AML) with normal cytogenetics.* PLoS One. **7**(8): p. e42334.
92. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc. **7**(3): p. 562-78.

93. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
94. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
95. Kim, D. and S.L. Salzberg, *TopHat-Fusion: an algorithm for discovery of novel fusion transcripts.* Genome Biol. **12**(8): p. R72.
96. Melo, J.V. and D.J. Barnes, *Chronic myeloid leukaemia as a model of disease evolution in human cancer.* Nat Rev Cancer, 2007. **7**(6): p. 441-53.
97. Calabretta, B. and D. Perrotti, *The biology of CML blast crisis.* Blood, 2004. **103**(11): p. 4010-22.
98. Ernst, T. and A. Hochhaus, *Chronic myeloid leukemia: clinical impact of BCR-ABL1 mutations and other lesions associated with disease progression.* Semin Oncol, 2012. **39**(1): p. 58-66.
99. Grossmann, V., et al., *A deep-sequencing study of chronic myeloid leukemia patients in blast crisis (BC-CML) detects mutations in 76.9% of cases.* Leukemia, 2011. **25**(3): p. 557-60.
100. Feinstein, E., et al., *p53 in chronic myelogenous leukemia in acute phase.* Proc Natl Acad Sci U S A, 1991. **88**(14): p. 6293-7.
101. Roche-Lestienne, C., et al., *RUNX1 DNA-binding mutations and RUNX1-PRDM16 cryptic fusions in BCR-ABL+ leukemias are frequently associated with secondary trisomy 21 and may contribute to clonal evolution and imatinib resistance.* Blood, 2008. **111**(7): p. 3735-41.
102. Sill, H., J.M. Goldman, and N.C. Cross, *Homozygous deletions of the p16 tumor-suppressor gene are associated with lymphoid transformation of chronic myeloid leukemia.* Blood, 1995. **85**(8): p. 2013-6.
103. Mullighan, C.G., et al., *BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros.* Nature, 2008. **453**(7191): p. 110-4.
104. Johansson, B., T. Fioretos, and F. Mitelman, *Cytogenetic and molecular genetic evolution of chronic myeloid leukemia.* Acta Haematol, 2002. **107**(2): p. 76-94.
105. Hochhaus, A. and P. La Rosee, *Imatinib therapy in chronic myelogenous leukemia: strategies to avoid and overcome resistance.* Leukemia, 2004. **18**(8): p. 1321-31.
106. O'Hare, T., C.A. Eide, and M.W. Deininger, *Bcr-Abl kinase domain mutations, drug resistance, and the road to a cure for chronic myeloid leukemia.* Blood, 2007. **110**(7): p. 2242-9.
107. Talpaz, M., et al., *Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias.* N Engl J Med, 2006. **354**(24): p. 2531-41.
108. Kantarjian, H., et al., *Nilotinib in imatinib-resistant CML and Philadelphia chromosome-positive ALL.* N Engl J Med, 2006. **354**(24): p. 2542-51.
109. Boultwood, J., et al., *High-density single nucleotide polymorphism array analysis and ASXL1 gene mutation screening in chronic myeloid leukemia during disease progression.* Leukemia. **24**(6): p. 1139-45.
110. Makishima, H., et al., *CBL, CBLB, TET2, ASXL1, and IDH1/2 mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia.* Blood. **117**(21): p. e198-206.
111. Piccaluga, P.P., et al., *Cytoplasmic mutated nucleophosmin (NPM1) in blast crisis of chronic myeloid leukaemia.* Leukemia, 2009. **23**(7): p. 1370-1.

112. Boultwood, J., et al., *High-density single nucleotide polymorphism array analysis and ASXL1 gene mutation screening in chronic myeloid leukemia during disease progression.* Leukemia, 2010. **24**(6): p. 1139-45.
113. Soverini, S., et al., *IDH2 somatic mutations in chronic myeloid leukemia patients in blast crisis.* Leukemia, 2011. **25**(1): p. 178-81.
114. Watkins, D.B., et al., *NPM1 mutations occur rarely or not at all in chronic myeloid leukaemia patients in chronic phase or blast crisis.* Leukemia, 2013. **27**(2): p. 489-90.
115. Forbes, S.A., et al., *COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer.* Nucleic Acids Res. **39**(Database issue): p. D945-50.
116. Makishima, H., et al., *CBL, CBLB, TET2, ASXL1, and IDH1/2 mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia.* Blood, 2011. **117**(21): p. e198-206.
117. Larsson NE, L.H., Rissler M, Högberg C, Richter J, Fioretos T, *et al*, *Exome sequencing during disease evolution in CML patients displaying an isochromosome 17q*. 2013, *Eur J Hum Genet* Paris. p. 285.
118. Tefferi, A., *Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: JAK2, MPL, TET2, ASXL1, CBL, IDH and IKZF1.* Leukemia, 2010. **24**(6): p. 1128-38.
119. Jager R, G.H., Berg T, Passamonti F, Cazzola M, Rumi E *et al*. , *Deletions of the Transcription Factor Ikaros in Myeloproliferative Neoplasms at Transformation to Acute Myeloid Leukemia*, ASH, Editor. 2012, Blood: Atlanta. p. 435.
120. Metcalf, D., *The granulocyte-macrophage colony stimulating factors.* Cell, 1985. **43**(1): p. 5-6.
121. Pardanani, A., et al., *CSF3R T618I is a highly prevalent and specific mutation in chronic neutrophilic leukemia.* Leukemia.
122. Kosmider, O., et al., *Mutation of the colony-stimulating factor-3 receptor gene is a rare event with poor prognosis in chronic myelomonocytic leukemia.* Leukemia, 2013.
123. Makishima, H., et al., *Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis.* Blood, 2012. **119**(14): p. 3203-10.
124. Jerez, A., et al., *Loss of heterozygosity in 7q myeloid disorders: clinical associations and genomic pathogenesis.* Blood. **119**(25): p. 6109-17.
125. Makishima, H., et al., *Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis.* Blood. **119**(14): p. 3203-10.
126. Wu, Y., et al., *Accelerated hepatocellular carcinoma development in mice expressing the Pim-3 transgene selectively in the liver.* Oncogene. **29**(15): p. 2228-37.
127. Forshell, L.P., et al., *The direct Myc target Pim3 cooperates with other Pim kinases in supporting viability of Myc-induced B-cell lymphomas.* Oncotarget. **2**(6): p. 448-60.
128. Gotlib, J., et al., *The new genetics of chronic neutrophilic leukemia and atypical CML: implications for diagnosis and treatment.* Blood.
129. James, L.I. and S.V. Frye, *Targeting chromatin readers.* Clin Pharmacol Ther, 2013. **93**(4): p. 312-4.

130. Burridge, S., *Target watch: Drugging the epigenome.* Nat Rev Drug Discov, 2013. **12**(2): p. 92-3.
131. Abdel-Wahab, O., et al., *Genetic analysis of transforming events that convert chronic myeloproliferative neoplasms to leukemias.* Cancer Res, 2010. **70**(2): p. 447-52.
132. Tefferi, A., *Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: JAK2, MPL, TET2, ASXL1, CBL, IDH and IKZF1.* Leukemia. **24**(6): p. 1128-38.
133. Holmfeldt, L., et al., *The genomic landscape of hypodiploid acute lymphoblastic leukemia.* Nat Genet. **45**(3): p. 242-52.
134. Caudell, D., et al., *Retroviral insertional mutagenesis identifies Zeb2 activation as a novel leukemogenic collaborating event in CALM-AF10 transgenic mice.* Blood. **115**(6): p. 1194-203.
135. Shih, A.H., et al., *The role of mutations in epigenetic regulators in myeloid malignancies.* Nat Rev Cancer. **12**(9): p. 599-612.
136. Metzeler, K.H., et al., *TET2 mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: a Cancer and Leukemia Group B study.* J Clin Oncol. **29**(10): p. 1373-81.
137. Bejar, R., et al., *Clinical effect of point mutations in myelodysplastic syndromes.* N Engl J Med. **364**(26): p. 2496-506.
138. Tefferi, A., et al., *TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis.* Leukemia, 2009. **23**(5): p. 905-11.
139. Green, C.L., et al., *The prognostic significance of IDH1 mutations in younger adult patients with acute myeloid leukemia is dependent on FLT3/ITD status.* Blood. **116**(15): p. 2779-82.
140. Green, C.L., et al., *The prognostic significance of IDH2 mutations in AML depends on the location of the mutation.* Blood. **118**(2): p. 409-12.
141. Tefferi, A., et al., *IDH mutations in primary myelofibrosis predict leukemic transformation and shortened survival: clinical evidence for leukemogenic collaboration with JAK2V617F.* Leukemia. **26**(3): p. 475-80.
142. Walter, M.J., et al., *Recurrent DNMT3A mutations in patients with myelodysplastic syndromes.* Leukemia. **25**(7): p. 1153-8.

CHAPTER 7: Publications

1. **JULIANE MENEZES**, Hideki Makishima, Ines Gomez, Francesco Acquadro, Gonzalo Gómez-López, Osvaldo Graña, Ana Dopazo, Sara Álvarez, Mercedes Trujillo, David G. Pisano, Jaroslaw P. Maciejewski and Juan C. Cigudosa. *CSF3R* T618I co-occurs with mutations of splicing and epigenetic genes and with a newly *PIM3* truncated fusion gene in chronic neutrophilic leukemia. *Submitted.*

2. **JULIANE MENEZES***, Rocío N. Salgado*, Francesco Acquadro, Gonzalo Gómez-López, Maria Carmen Carralero, Alicia Barroso, Fatima Mercadillo, Luis Espinosa-Hevia, Juan Garcia Talavera-Casañas, David G. Pisano, Sara Álvarez and Juan C. Cigudosa (*Both authors contributed equally to this work). Mutations in *ASXL1*, *IKZF3* and *TP53* are players in chronic myeloid leukemia initiation and progression. *Submitted.*

3. **JULIANE MENEZES**, Francesco Acquadro, Matthew Wiseman, Gonzalo Gómez-López, Rocío N. Salgado, Juan Garcia Talavera-Casañas, Ismael Buño, José V. Cervera, Santiago Montes-Moreno, Jesús María Hernández-Rivas, Rosa Ayala, Maria José Calasanz, Maria José Larrayoz, Lourdes Florensa Brichs, Marta Gonzalez-Vicent, David G. Pisano, Miguel Angel Piris, Sara Álvarez and Juan C. Cigudosa. Exome sequencing reveals novel and recurrent mutations with clinical impact in Blastic Plasmacytoid Dendritic Cell Neoplasm. *Submitted.*

4. Rocío N. Salgado*, **JULIANE MENEZES***, Maria Calvente, Javier Suela, Francesco Acquadro, Carolina Martínez-Laperche, Rafael Flores, Mercedes Trujillo, Sara Álvarez and Juan C. Cigudosa (*Both authors contributed equally to this work). Myeloid neoplasms with der(1)t(1;19) are characterized by cooperating gene mutations specially involved in epigenetics. *Submitted.*

5. Elena Doménech, Gonzalo Gómez-López, Daniel Gzlez-Peña, Mar López, Beatriz Herreros, **JULIANE MENEZES**, Natalia Gómez-Lozano, Angel Carro, Osvaldo Graña, David G. Pisano, Orlando Domínguez, José A. García-Marco, Miguel Angel Piris and Margarita Sánchez-Beato. New mutations in chronic lymphocytic leukemia identified by target enrichment and deep sequencing. *PLoS One*. 2012; 7(6):e38158.

6. Ana del Rio-Machín*, **JULIANE MENEZES***, Alba Maiques-Diaz, Bibiana Ferreira, Francesco Acquadro, Sandra Rodriguez-Perales, Karmele Arribalzaga Juaristi, Sara Álvarez and Juan C. Cigudosa. (*Both authors contributed equally to this work). Abrogation of *RUNX1* gene expression in *"de novo"* myelodysplastic syndrome with t(4;21)(q21;q22). *Haematologica*. 2012 Apr; 97(4):534-7.

7. **JULIANE MENEZES**, Francesco Acquadro, Concepción Perez-Pons de la Villa, Félix García-Sánchez, Sara Álvarez and Juan C. Cigudosa. FIP1L1/RARA with breakpoint at *FIP1L1* intron 13: a variant translocation in acute promyelocytic leukaemia. *Haematologica*. 2011; 96: 1565-6.

CHAPTER 8: Molecular Cytogenetics Group

Molecular Cytogenetics Group 2009

Group Leader: Juan C. Cigudosa
Staff Scientists: Sara Álvarez, Sandra Rodríguez
Graduate Students: Bibiana I. Ferreira, Ana del Río, Alba Maiques, Juliane Menezes, Jaroslaw K. Sochacki.
Technicians: Francesco Acquadro, Carmen Carralero, Miguel A. Grillo, M. Carmen Martín, Gloria Soler

Molecular Cytogenetics Group 2010

Group Leader: Juan C. Cigudosa
Staff Scientists: Sara Álvarez, Sandra Rodríguez
Graduate Students: Ana del Río, Alba Maiques, Juliane Menezes, Jaroslaw K. Sochacki
Technicians: Francesco Acquadro, Carmen Carralero, Almudena Gil, Miguel A. Grillo, M. Carmen Martín

Molecular Cytogenetics Group 2011

Group Leader: Juan C. Cigudosa
Staff Scientists: Sara Álvarez, Sandra Rodríguez, Margarita Sánchez-Beato
Graduate Students: Ana del Río, Alba Maiques, Juliane Menezes, Jaroslaw K. Sochacki
Technicians: Francesco Acquadro, Carmen Carralero, Almudena Gil, Luis Espinosa, Miguel A. Grillo, M. Carmen Martín

Molecular Cytogenetics Group 2012

Group Leader: Juan C. Cigudosa
Staff Scientists: Sara Álvarez, Sandra Rodríguez
Graduate Students: Carlos Benítez, Ana del Río, Alba Maiques, Juliane Menezes, Jaroslaw K. Sochacki
Technicians: Francesco Acquadro, Carmen Carralero, Luis Espinosa, Miguel A. Grillo, Miriam Hernando, M. Carmen Martín, Rocío Nieves Salgado